Centralized Versus Decentralized Team Games of Distributed Stochastic Differential Decision Systems with Noiseless Information Structures-Part I: General Theory

Charalambos D. Charalambous and Nasir U. Ahmed

Abstract

Decentralized optimization of distributed stochastic differential systems has been an active area of research for over half a century. Its formulation utilizing static team and person-by-person optimality criteria is well investigated. However, the results have not been generalized to nonlinear distributed stochastic differential systems possibly due to technical difficulties inherent with decentralized decision strategies.

In this first part of the two-part paper, we derive team optimality and person-by-person optimality conditions for distributed stochastic differential systems with different information structures. The optimality conditions are given in terms of a Hamiltonian system of equations described by a system of coupled backward and forward stochastic differential equations and a conditional Hamiltonian, under both regular and relaxed strategies. Our methodology is based on the semi martingale representation theorem and variational methods. Throughout the presentation we discuss similarities to optimality conditions of centralized decision making.

Index Terms. Team and Person-by-Person Optimality, Stochastic Differential Systems, Stochastic Maximum Principle, Relaxed Strategies.

C.D. Charalambous is with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678 (E-mail: chadcha@ucy.ac.cy).

N.U Ahmed is with the School of Engineering and Computer Science, and Department of Mathematics, University of Ottawa, Ontario, Canada, K1N 6N5 (E-mail: ahmed@site.uottawa.ca).

I. INTRODUCTION

Over the last 50 years many mathematical concepts and procedures were developed to design optimal control strategies for stochastic dynamical systems. We refer to this set of mathematical concepts and procedures as the ”classical theory of stochastic optimization”. It has been utilized extensively to address the questions of existence of optimal strategies, and necessary and sufficient optimality conditions for systems driven by continuous martingale processes (Brownian motion processes), and discontinuous martingale processes (jump processes). It has been successfully applied to centralized fully observable control problems, meaning the admissible strategies are functions of a common noiseless measurements of the system [1]–[9], and to centralized partially observable control systems, meaning the admissible strategies are functions of common noisy measurements of the system [2], [10]–[13]. In addition, optimility conditions are derived for infinite dimensional systems and impulsive systems in [4], [9], [14]. Thus, the classical theory of optimization is developed on the assumption of centralized decisions or control actions. It presupposes that all information about the system can be acquired and accordingly the decision policies (control actions) can be formulated. The basic underlying assumption is that the acquisition of the information is centralized or the information acquired at different locations is communicated to each decision maker or control.

When the system model consists of multiple decision makers, and the acquisition of information and its processing is decentralized or shared among several locations, the decision makers actions are based on different information. We call the information available for such decisions, ”decentralized information structures or patterns”. When the system model is dynamic, consisting of an interconnection of at least two subsystems, and the decisions are based on decentralized information structures, we call the overall system a ”distributed system with decentralized information structures”. Over the years several specific forms of decentralized information structures are analyzed mostly in discrete-time [15]–[26], and more recently [27]–[32]. However, at this stage there is no systematic framework addressing optimality conditions for distributed systems with decentralized information structures. The absence of such optimization theory raises the question whether the classical theory of optimization is limited in mathematical concepts and procedures to deal with decentralized systems.

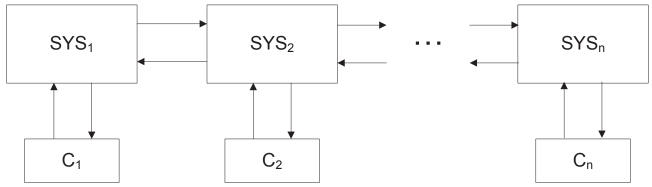

Fig. 1. Diagram of architecture for distributed stochastic differential decision systems.

In this first part of the two-part investigation, we show that the classical theory of optimization does not have such a limitation. We consider a team game reward [23], [26], [33]–[35] and we apply concepts from the classical theory of optimization to derive necessary and sufficient optimality conditions for nonlinear stochastic distributed systems with decentralized information structures. Our methodology utilizes the semi martingale representation theorem and variational methods recently reported by the authors in [36].

The optimality conditions developed in this paper can be applied to many architectures of distributed systems such as Fig. 1 (see also [37]). Each decision maker makes its decision based on local information and exerts control action that affects the overall distributed system, without allowing communication between the local decision makers. Such systems are called distributed systems with decentralized information structures. The team formulation of the distributed system with decentralized information structures, consists of an interconnection of N subsystems. Each subsystem i has its state denoted by $x^i \in \mathbb{X}^i$, a local decision maker or control input $u^i \in \mathbb{A}^i$, an exogenous Brownian motion noise input $W^i \in \mathbb{W}^i$, and a coupling from the other subsystem.

Decentralized Information Structures for Decision Makers

The information structures of the local decision makers $u^i, i = 1, 2, \ldots, N$ are defined as follows. For any $t \in [0, T]$, the information structure available to decision maker (DM) u^i is modeled by the $\sigma-$algebra $\mathcal{G}^i_{0,t}$ generated by the observable events associated with the local subsystem. These observables can be generated by nonanticipative functionals of the noise entering the system, nonanticipative functions of the state of the system, its delayed versions, or any possible combinations thereof. Let us denote the admissible strategies of u^i with action spaces

$\mathbb{A}^i$, by $\mathbb{U}^i[0,T], i = 1,2,\ldots,N$ (meaning that u^i is a nonanticipative measurable functional of the information algebra $\mathcal{G}^i_T \stackrel{\triangle}{=} \{\mathcal{G}^i_{0,t} : t \in [0,T]\}$ taking values from $\mathbb{A}^i$. Thus the augmented state, control and noise of the decentralized system can be written as

$$x \stackrel{\triangle}{=} (x^1, x^2, \ldots, x^N) \in \mathbb{X}^{(N)}, \quad u \stackrel{\triangle}{=} (u^1, u^2, \ldots, u^N) \in \mathbb{A}^{(N)}, \quad W \stackrel{\triangle}{=} (W^1, W^2, \ldots, W^N) \in \mathbb{W}^{(N)}.$$

Then the overall system can be expressed in compact form by the following stochastic Itô differential equation

$$dx(t) = f(t, x(t), u_t)dt + \sigma(t, x_t, u_t)dW(t), \quad x(0) = x_0, \quad t \in (0,T]. \tag{1}$$

Team Game Pay-off Functional

The objective is to find a team optimal strategy $u^o \equiv (u^{1,o}, \ldots, u^{N,o}) \in \times_{i=1}^N \mathbb{U}^i[0,T]$ at which the pay-off functional defined by

$$J(u^o) \equiv J(u^{1,o}, \ldots, u^{N,o}) \stackrel{\triangle}{=} \inf_{(u^1 \ldots u^N) \in \times_{i=1}^N \mathbb{U}^i[0,T]} \mathbb{E}\Big\{ \int_0^T \ell(t, x(t), u(t))dt + \varphi(x(T)) \Big\} \tag{2}$$

attains its minimum.

We consider two main classes of decentralized noiseless information structures; 1) nonanticipative functionals of any subset of the sybsystems Brownian motions $\{W^1, \ldots, W^N\}$, called ”nonanticipative information structures”, and 2) nonanticipative functionals of any subset of the subsystem states $\{x^1, \ldots, x^N\}$, called ”feedback information structures” (see Section II-C).

Team Game Optimality Conditions

In Section V we derive team optimality conditions (Theorem 9) for pay-off (2) subject to (1), under a strong formulation of the filtered probability space $\Big(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P}\Big)$. These are summarized below.

Define the Hamiltonian

$$\mathcal{H} : [0,T] \times \mathbb{X}^{(N)} \times \mathbb{X}^{(N)} \times \mathcal{L}(\mathbb{W}^{(N)}, \mathbb{X}^{N)}) \times \mathbb{A}^{(N)} \longrightarrow \mathbb{R}$$

by

$$\mathcal{H}(t, \xi, \zeta, M, \nu) \stackrel{\triangle}{=} \langle f(t, \xi, \nu), \zeta \rangle + tr(M^* \sigma(t, \xi, \nu)) + \ell(t, \xi, \nu), \quad t \in [0,T]. \tag{3}$$

For any $u \in \mathbb{U}^{(N)} \equiv \times_{i=1}^{N} \mathbb{U}^i[0,T]$, consider the adjoint process $\{\psi, Q\}$ and the state x satisfying the following backward and forward stochastic differential equations respectively,

$$d\psi(t) = -\mathcal{H}_x(t, x(t), \psi(t), Q(t), u_t)dt + Q(t)dW(t), \quad \psi(T) = \varphi_x(x(T)), \quad t \in [0,T), \tag{4}$$

$$dx(t) = \mathcal{H}_\psi(t, x(t), \psi(t), Q(t), u_t)dt + \sigma(t, x(t), u_t)dW(t), \quad x(0) = x_0, \quad t \in (0,T]. \tag{5}$$

The stochastic optimality conditions of the team game with decentralized noiseless information structures are given below.

(1) **Necessary Conditions.** Under certain conditions, which are precisely those of the classical theory of optimization, the following hold.
For an element $u^o \in \mathbb{U}^{(N)} \equiv \times_{i=1}^{N} \mathbb{U}^i[0,T]$ with the corresponding solution x^o to be team optimal, it is necessary that the following hold:
The process $\{\psi^o, Q^o\}$ is the unique solution of the backward stochastic differential equation (4) corresponding to the pair $\{u^o, x^o\}$ and that they together satisfy the point wise almost sure inequalities with respect to the σ-algebras $\mathcal{G}^i_{0,t}$, $t \in [0,T], i = 1, 2, \ldots, N$:

$$\begin{aligned}
&\mathbb{E}\Big\{\mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{1,o}, \ldots, u_t^{i-1,o}, u^i, u_t^{i+1,o}, \ldots, u_t^{N,o})|\mathcal{G}^i_{0,t}\Big\} \\
&\quad \geq \mathbb{E}\Big\{\mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{1,o}, \ldots, u_t^{i-1,o}, u_t^{i,o}, u_t^{i+1,o}, \ldots, u_t^{N,o})|\mathcal{G}^i_{0,t}\Big\}, \\
&\quad\quad \forall u^i \in \mathbb{A}^i, \quad a.e. t \in [0,T], \quad \mathbb{P}|_{\mathcal{G}^i_{0,t}} - a.s., \quad i = 1, 2, \ldots, N.
\end{aligned} \tag{6}$$

(2) **Sufficient Conditions.** Under global convexity of the Hamiltonian with respect to the state and control variables and convexity of the terminal pay-off function $\varphi(\cdot)$ the pair $\{x^o(\cdot), u^o(\cdot)\}$ is optimal if it satisfies (6).

An important feature obtained during the derivation is that the optimality conditions for a team optimal strategy are equivalent to the optimality conditions for a person-by-person optimal strategy. This follows from Theorem 6 and Corollary 1.

The point to be made regarding the derivation of the above optimality conditions, is that we convert the problem into a centralized problem with the associated Hamiltonian system of equations to capture the constraints, and only at the final step, the optimality of decentralized strategies is addressed, by identifying the conditional variational Hamiltonian which is consistent with the decentralized information structures. That is, the Hamiltonian system (4), (5) is the one

corresponding to centralized strategies, while the conditional Hamiltonian (6) is the projection of the centralized Hamiltonian onto the subspace generated by the decentralized information structures.

We conclude the preliminary discussion on classical optimization theory of centralized strategies versus decentralized strategies, by stating that there are no limitations in applying classical theory of optimization to distributed systems with decentralized information structures. Rather, the challenge is in the computation of the conditional Hamiltonians, and hence the optimal strategies. However, this has also remained a challenge for centralized fully or partially observed strategies.

The specific objectives of this paper are the following.

(a) Derive team games necessary conditions of optimality (stochastic maximum principle) for distributed stochastic differential systems with decentralized information structures.

(b) Introduce assumptions so that the team games necessary conditions of optimality in **(a)** are also sufficient;

(c) Derive person-by-person optimality conditions and discuss their relation with team optimality conditions;

(d) Prove existence of optimal team and person-by-person strategies for distributed stochastic differential systems with decentralized information structures, using the theory of relaxed control strategies, and relate **(a), (b), (c)** to regular decision strategies.

A detailed investigation of applications of the results of this part to specific linear and nonlinear distributed stochastic differential decision systems is discussed in the second part of this two-part paper [38] where we derive the explicit expressions for the optimal decentralized strategies.

The rest of the paper is organized as follows. In Section II we formulate the distributed stochastic differential system with decentralized information structures. In Section III, we consider the question of existence of optimal relaxed controls (decisions). In Section IV, we develop the stochastic optimality conditions for team games with decentralized information structures, consisting of necessary and sufficient conditions of optimality. In Section V, we specialize the necessary and sufficient optimality conditions to regular strategies and obtain corresponding

necessary and sufficient optimality conditions. The paper is concluded with some comments on possible extensions of our results.

II. Team Games of Stochastic Differential Systems

In this section we introduce the mathematical formulation of distributed stochastic systems, the information structures available to the decision makers for their actions, and the definitions of collaborative decisions via team game optimality and person-by-person optimality. Throughout the terms "decision maker" or "control" are used interchangeably. A stochastic dynamical decision or control system is called distributed if it consists of an interconnection of at least two subsystems and decision makers. The underlying assumption for these distributed systems is that the decision makers actions are based on decentralized information structures. However, the decision makers are allowed to exchange information on their law or strategy deployed, e.g., the functional form of their strategies but not their actions.

Some Basic Terminologies

DM	Abbreviation for "Decision Maker"
$\mathbb{Z}_N \stackrel{\triangle}{=} \{1, 2, \ldots, N\}$	subset of natural numbers
$s \stackrel{\triangle}{=} \{s^1, s^2, \ldots, \ldots, s^N\}$	set consisting of N elements
$s^{-i} = s \setminus \{s^i\}, \ \ s = (s^{-i}, s^i)$	set s minus $\{s^i\}$
$\mathcal{L}(\mathcal{X}, \mathcal{Y})$	linear transformation mapping a vector space $\mathcal{X}$ into a vector space $\mathcal{Y}$
$A^{(i)}$	ith column of a map $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $i = 1, \ldots, n$
$(\mathbb{A}^i, d)$	separable metric space for player $i \in \mathbb{Z}_N$ actions
$\mathbb{A}^{(N)} \stackrel{\triangle}{=} \times_{i=1}^N \mathbb{A}^i$	product action space of N players
$\mathbb{U}^i_{reg}[0, T]$	regular admissible strategy of player $i \in \mathbb{Z}_N$
$\mathbb{U}^i_{rel}[0, T]$	relaxed admissible strategy of player $i \in \mathbb{Z}_N$

Let $\Big(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P}\Big)$ denote a complete filtered probability space satisfying the usual conditions [39], that is, $(\Omega, \mathbb{F}, \mathbb{P})$ is complete, $\mathbb{F}_{0,0}$ contains all $\mathbb{P}$-null sets in $\mathbb{F}$. Note that filtrations $\{\mathbb{F}_{0,t} : t \in [0,T]\}$ are monotone in the sense that $\mathbb{F}_{0,s} \subseteq \mathbb{F}_{0,t}$, $\forall 0 \leq s \leq t \leq T$. Moreover, $\{\mathbb{F}_{0,t} : t \in [0,T]\}$ is called right continuous if $\mathbb{F}_{0,t} = \mathbb{F}_{0,t+} \stackrel{\triangle}{=} \bigcap_{s>t} \mathbb{F}_{0,s}, \forall t \in [0,T)$ and it is called left continuous if $\mathbb{F}_{0,t} = \mathbb{F}_{0,t-} \stackrel{\triangle}{=} \sigma\Big(\bigcup_{s<t} \mathbb{F}_{0,s}\Big), \forall t \in (0,T]$. Throughout the paper filtrations are denoted by $\mathbb{F}_T \stackrel{\triangle}{=} \{\mathbb{F}_{0,t} : t \in [0,T]\}$, and they are assumed to be right continuous and complete.

Consider a random process $\{z(t) : t \in [0,T]\}$ defined on the filtered probability space $(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P})$ and taking values in a metric space $(\mathbb{Z}, d)$. The process $\{z(t) : t \in [0,T]\}$ is said to be measurable if the map $(t,\omega) \rightarrow z(t,\omega)$ is $\mathcal{B}([0,T]) \times \mathbb{F}/\mathcal{B}(\mathbb{Z})-$measurable where $\mathcal{B}(\mathbb{Z})$ denotes the Borel algebra of subsets of $\mathbb{Z}$. The process $\{z(t) : t \in [0,T]\}$ is said to be $\{\mathbb{F}_{0,t} : t \in [0,T]\}-$adapted if for all $t \in [0,T]$, the map $\omega \rightarrow z(t,\omega)$ is $\mathbb{F}_{0,t}/\mathcal{B}(\mathbb{Z})-$measurable. The process $\{z(t) : t \in [0,T]\}$ is said to be $\{\mathbb{F}_{0,t} : t \in [0,T]\}-$progresively measurable if for all $t \in [0,T]$, the map $(s,\omega) \rightarrow z(s,\omega)$ is $\mathcal{B}([0,t]) \otimes \mathbb{F}_{0,t}/\mathcal{B}(\mathbb{Z})-$measurable. It can be shown that any stochastic process $\{z(t) : t \in [0,T]\}$ on a filtered probability space $(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P})$ which is measurable and adapted has a progressively measurable modification [39]. Unless otherwise specified, we shall say a process $\{z(t) : t \in [0,T]\}$ is $\{\mathbb{F}_{0,t} : t \in [0,T]\}-$adapted if the processes is $\{\mathbb{F}_{0,t} : t \in [0,T]\}-$progressively measurable.

In our derivations we make extensive use of the following spaces considered by the authors in [36]. Let $L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n) \subset L^2(\Omega \times [0,T], d\mathbb{P} \times dt, \mathbb{R}^n) \equiv L^2([0,T], L^2(\Omega, \mathbb{R}^n))$ denote the space of $\mathbb{F}_T-$adapted random processes $\{z(t) : t \in [0,T]\}$ such that

$$\mathbb{E} \int_{[0,T]} |z(t)|^2_{\mathbb{R}^n} dt < \infty,$$

which is a sub-Hilbert space of $L^2([0,T], L^2(\Omega, \mathbb{R}^n))$. Similarly, let $L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)) \subset L^2([0,T], L^2(\Omega, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)))$ denote the space of $\mathbb{F}_T-$adapted $n \times m$ matrix valued random processes $\{\Sigma(t) : t \in [0,T]\}$ such that

$$\mathbb{E} \int_{[0,T]} |\Sigma(t)|^2_{\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)} dt \stackrel{\triangle}{=} \mathbb{E} \int_{[0,T]} tr(\Sigma^*(t)\Sigma(t)) dt < \infty.$$

A. Regular Strategies

In this subsection we consider measurable vector valued functions, also known as regular strategies. We consider the strong formulation. Let $\Big(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P}\Big)$ denote a fixed complete filtered probability space on which are based all random processes considered in the paper. At this stage we do not specify how $\{\mathbb{F}_{0,t} : t \in [0,T]\}$ came about, but we require that Brownian motions are adapted to this filtration.

Admissible Decision Maker Strategies

The Decision Makers (DM) $\{u^i : i \in \mathbb{Z}_N\}$ take values in a closed convex subset of linear metric spaces $\{(\mathbb{M}^i, d) : i \in \mathbb{Z}_N\}$. Let $\mathcal{G}_T^i \stackrel{\triangle}{=} \{\mathcal{G}_{0,t}^i : t \in [0,T]\} \subset \{\mathbb{F}_{0,t} : t \in [0,T]\}$ denote the information available to DM i, $\forall i \in \mathbb{Z}_N$. The admissible set of regular strategies is defined by

$$\mathbb{U}_{reg}^i[0,T] \stackrel{\triangle}{=} \Big\{ u^i \in L^2_{\mathcal{G}_T^i}([0,T], \mathbb{R}^{d_i}) : \; u_t^i \in \mathbb{A}^i \subset \mathbb{R}^{d_i}, \; a.e.t \in [0,T], \; \mathbb{P}-a.s. \Big\}, \quad \forall i \in \mathbb{Z}_N. \quad (7)$$

Clearly, $\mathbb{U}_{reg}^i[0,T]$ is a closed convex subset of $L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^{d_i})$, for $i = 1, 2, \ldots, N$. That is, $u^i : [0,T] \times \Omega \rightarrow \mathbb{A}^i$, and $\{u_t^i : t \in [0,T]\}$ is $\mathcal{G}_T^i-$adapted, $\forall i \in \mathbb{Z}_N$.
An N tuple of DM strategies is by definition $(u^1, u^2, \ldots, u^N) \in \mathbb{U}_{reg}^{(N)}[0,T] \stackrel{\triangle}{=} \times_{i=1}^N \mathbb{U}_{reg}^i[0,T]$, which are nonanticipative with respect to the information structures $\{\mathcal{G}_{0,t}^i : t \in [0,T]\}, i = 1, 2, \ldots, N$. Hence, the information structure of each DM, $\mathcal{G}_T^i$, is decentralized, and may be generated by local or global subsystem observables. Nonanticipative strategies are often utilized when deriving the minimum principle for centralized stochastic control or decision systems [8].

Distributed Stochastic Systems

Given a fixed probability space $\Big(\Omega, \mathbb{F}, \{\mathbb{F}_{0,t} : t \in [0,T]\}, \mathbb{P}\Big)$, a distributed stochastic system consists of an interconnection of N subsystems. Each subsystem i has its own state space $\mathbb{R}^{n_i}$, action space $\mathbb{A}^i \subset \mathbb{R}^{d_i}$, an exogenous noise space $\mathbb{W}^i \stackrel{\triangle}{=} \mathbb{R}^{m_i}$, and an initial state $x^i(0) = x_0^i$, identified by the following quantities.

(S1) $x^i(0) = x_0^i$: an $\mathbb{R}^{n_i}$-valued Random Variable;

(S2) $\{W^i(t) : t \in [0,T]\}$: an $\mathbb{R}^{m_i}$-valued standard Brownian motion which models the exogenous state noise, adapted to $\mathbb{F}_T$, independent of $x^i(0)$.

Each subsystem is described by a finite dimensional system of coupled stochastic differential equations of Itô type as follows.

$$\begin{aligned} dx^i(t) =& f^i(t,x^i(t),u^i_t)dt + \sigma^i(t,x^i(t),u^i_t)dW^i(t) + \sum_{j=1,j\neq i}^N f^{ij}(t,x^j(t),u^j_t)dt \\ &+ \sum_{j=1,j\neq i}^N \sigma^{ij}(t,x^j(t),u^j_t)dW^j(t), \quad x^i(0)=x^i_0, \quad t\in(0,T], \quad i\in\mathbb{Z}_N. \end{aligned} \tag{8}$$

On the product space $(\mathbb{X}^{(N)}, \mathbb{A}^{(N)}, \mathbb{W}^{(N)})$, where $\mathbb{X}^{(N)} \triangleq \times_{i=1}^N \mathbb{R}^{n_i}, \mathbb{A}^{(N)} \triangleq \times_{i=1}^N \mathbb{A}^i, \mathbb{W}^{(N)} \triangleq \times_{i=1}^N \mathbb{R}^{m_i}$, one defines the augmented vectors by

$$W \triangleq (W^1, W^2, \ldots, W^N) \in \mathbb{R}^m, \quad u \triangleq (u^1,u^2,\ldots,u^N) \in \mathbb{R}^d, \quad x \triangleq (x^1,x^2,\ldots,x^N)\in\mathbb{R}^n.$$

Then on the product space the distributed system is described in compact form by

$$dx(t) = f(t,x(t),u_t)dt + \sigma(t,x(t),u_t)\; dW(t), \quad x(0)=x_0, \quad t\in(0,T], \tag{9}$$

where $f:[0,T]\times\mathbb{R}^n\times\mathbb{A}^{(N)} \longrightarrow \mathbb{R}^n$ denotes the drift and $\sigma:[0,T]\times\mathbb{R}^n\times\mathbb{A}^{(N)} \longrightarrow \mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)$ the diffusion coefficients. Note that (9) is very general since no specific interconnection structure is assumed among the different subsystems.

Pay-off Functional

Consider the distributed system (9) with decentralized full information structures. Given a $u \in \mathbb{U}^{(N)}_{reg}[0,T]$, we define the reward or performance criterion by

$$J(u) \equiv J(u^1,u^2,\ldots,u^N) \triangleq \mathbb{E}\Big\{\int_0^T \ell(t,x(t),u_t)dt + \varphi(x(T))\Big\}, \tag{10}$$

where $\ell:[0,T]\times\mathbb{R}^n\times\mathbb{U}^{(N)} \longrightarrow (-\infty,\infty]$ denotes the integrand for the running cost functional and $\varphi:\mathbb{R}^n \longrightarrow (-\infty,\infty]$, the terminal cost function. Notice that the performance of the decentralized system is measured by a single pay-off functional. The interpretation is that there is a centralized layer where the quality of individual decision makers strategies are evaluated for a common goal. Therefore, the underlying assumption concerning the single pay-off instead of multiple pay-offs (one for each decision maker) is that the team objective can be met.

For deterministic as well as stochastic systems, it is well known that if the set $\mathbb{A}^i$ is not convex, there may not exist any optimal control. For this reason it is necessary to introduce relaxed strategies as discussed in the next subsection.

B. *Relaxed Strategies*

This paper will focus on relaxed strategies (also called randomized strategies) and later on specialize to regular strategies (measurable functions). Therefore, we introduce the formulation based on relaxed strategies (e.g. probability measures on the action space).

Distributed Stochastic Systems

For each $i \in \mathbb{Z}_N$, let $(\mathbb{M}^i, d)$ be a separable metric space with $\mathbb{A}^i \subset \mathbb{M}^i$ compact, and let $\mathcal{B}(\mathbb{A}^i)$ denote the Borel subsets of $\mathbb{A}^i$. Let $C(\mathbb{A}^i)$ denote the space of continuous functions on $\mathbb{A}^i$. Let $\mathcal{M}(\mathbb{A}^i)$ denote the space of regular bounded signed Borel measures on $\mathcal{B}(\mathbb{A}^i)$ and $\mathcal{M}_1(\mathbb{A}^i) \subset \mathcal{M}(\mathbb{A}^i)$ the space of regular probability measures. The DM strategies with different information structures on the time interval $[0,T]$ will be described through the topological dual of the Banach space $L^1_{\mathcal{G}^i_T}([0,T], C(\mathbb{A}^i))$, the L^1-space of $\mathcal{G}^i_T \stackrel{\triangle}{=} \{\mathcal{G}^i_{0,t} : t \in [0,T]]\}-$ adapted $C(\mathbb{A}^i)$ valued functions, for $i \in \mathbb{Z}_N$. For each $i \in \mathbb{Z}_N$ the dual of this space is given by $L^\infty_{\mathcal{G}^i_T}([0,T], \mathcal{M}(\mathbb{A}^i))$ which consists of weak* measurable $\mathcal{G}^i_T$ adapted $\mathcal{M}(\mathbb{A}^i)$ valued functions. The DM (control) strategies are drawn from the subspace $L^\infty_{\mathcal{G}^i_T}([0,T], \mathcal{M}_1(\mathbb{A}^i)) \subset L^\infty_{\mathcal{G}^i_T}([0,T], \mathcal{M}(\mathbb{A}^i))$. For convenience notation we denote this by

$$\mathbb{U}^i_{rel}[0,T] \stackrel{\triangle}{=} L^\infty_{\mathcal{G}^i_T}([0,T], \mathcal{M}_1(\mathbb{A}^i)), \quad i \in \mathbb{Z}_N, \tag{11}$$

and the team strategies by the product space

$$\mathbb{U}^{(N)}_{rel}[0,T] \stackrel{\triangle}{=} \times_{i=1}^N \mathbb{U}^i_{rel}[0,T], \quad \mathcal{M}_1(\mathbb{A}^N) \stackrel{\triangle}{=} \times_{i=1}^N \mathcal{M}_1(\mathbb{A}^i).$$

Thus, for any $i \in \mathbb{Z}_N$, given the information $\mathcal{G}^i_T$, player $\{u^i_t : t \in [0,T]\}$ is a stochastic kernel (conditional distribution) defined by

$$u^i_t(\Gamma) = q^i_t(\Gamma|\mathcal{G}^i_{0,t}), \quad \text{for } t \in [0,T], \text{ and } \quad \forall \Gamma \in \mathcal{B}(\mathbb{A}^i).$$

Clearly, for each $i \in \mathbb{Z}_N$ and for every $\varphi \in C(\mathbb{A}^i)$ the process

$$\int_{\mathbb{A}^i} \varphi(\xi) u^i_t(d\xi) = \int_{\mathbb{A}^i} \varphi(\xi) q^i_t(d\xi|\mathcal{G}^i_{0,t}), \quad t \in [0,T],$$

is $\mathcal{G}^i_T-$ progressively measurable. Given a $u \in \mathbb{U}^{(N)}_{rel}[0,T]$, the distributed system is written in compact form as

$$dx(t) = f(t, x(t), u_t)dt + \sigma(t, x(t), u_t)dW(t), \quad x(0) = x_0, \quad t \in [0,T], \tag{12}$$

where the drift and diffusion coefficient is now defined by

$$F(t,x,u_t) \stackrel{\triangle}{=} \int_{\mathbb{A}^{(N)}} \Big(b(t,x,\xi^1,\xi^2,\ldots,\xi^N)\Big) \times_{i=1}^N u_t^i(d\xi^i)dt, \quad t \in [0,T), \tag{13}$$

for $F = \{f, \sigma\}$,

Pay-off Functional

Given a $u \in \mathbb{U}_{rel}^{(N)}[0,T]$ the performance criterion is defined by

$$J(u) \stackrel{\triangle}{=} \mathbb{E}\bigg\{\int_0^T \ell(t,x(t),u_t)dt + \varphi(x(T))\bigg\} \tag{14}$$

$$\equiv \mathbb{E}\bigg\{\int_0^T \int_{\mathbb{A}^{(N)}} \Big(\ell(t,x(t),\xi^1,\xi^2,\ldots,\xi^N)\Big) \times_{i=1}^N u_t^i(d\xi^i)dt + \varphi(x(T))\bigg\} \tag{15}$$

where ℓ and φ are as defined before.

C. Team and Person-by-Person Optimality

In this section we give the precise definitions of team and person-by-person (i.e., player-by-player) optimality for relaxed and regular strategies. There are many possible information structures for control strategies $\{u^i : i \in \mathbb{Z}_N\}$. We consider the following.

(NIS): Nonanticipative Information Structures. Decision u^i is adapted to the filtration $\mathcal{G}_T^i \subset \mathbb{F}_T$ which is generated by the $\sigma-$ algebra induced by any combination of the subsystems Brownian motions and their increments $\{(W^1(t), W^2(t), \ldots, W^N(t)) : t \in [0,T]\}, \forall i \in \mathbb{Z}_N$. This is often called open loop information, and it is the one used in classical stochastic control with centralized full information to derive the maximum principe [8].

(FIS): Feedback Information Structures. Decision u^i is adapted to the filtration $\mathcal{G}_T^{z^i}$ generated by the $\sigma-$algebra $\mathcal{G}_{0,t}^{z^i} \stackrel{\triangle}{=} \sigma\{z^i(s) : 0 \leq s \leq t\}, t \in [0,T]$, where the observables z^i are nonanticipative measurable functionals of any combination of the states defined by

$$z^i(t) = h^i(t,x), \quad h^i : [0,T] \times C([0,T],\mathbb{R}^n) \longrightarrow \mathbb{R}^{k_i}, \quad i \in \mathbb{Z}_N. \tag{16}$$

Note that the state x and hence the observables z^i may depend on controls.
The set of admissible regular feedback strategies is defined by

$$\mathbb{U}_{reg}^{(N),z}[0,T] \stackrel{\triangle}{=} \Big\{ u \in \mathbb{U}_{reg}^{(N)}[0,T] : u_t^i \ \text{ is } \ \mathcal{G}_{0,t}^{z^i} - \text{measurable}, t \in [0,T], \quad i = 1, \ldots, N \Big\}. \tag{17}$$

Similarly, the set of admissible relaxed feedback strategies is defined by

$$\mathbb{U}_{rel}^{(N),z}[0,T] \stackrel{\triangle}{=} \Big\{ u \in \mathbb{U}_{rel}^{(N)}[0,T] : u^i \in L^\infty_{\mathcal{G}_T^{z^i}}([0,T], \mathcal{M}_1(\mathbb{A}^i)), \quad i = 1, \ldots, N \Big\}. \tag{18}$$

One might be tempted to believe that nonanticipative strategies might be restrictive, because they are not explicitly described in terms of feedback. We will show that this is not true. In fact such strategies cover a large number of interesting problems.

Problem 1. *(Team Optimality)*

(RS): Relaxed Strategies. *Given the pay-off functional (14), constraint (12) the N tuple of relaxed strategies $u^o \stackrel{\triangle}{=} (u^{1,o}, u^{2,o}, \ldots, u^{N,o}) \in \mathbb{U}_{rel}^{(N)}[0,T]$ is called nonanticipative team optimal if it satisfies*

$$J(u^{1,o}, u^{2,o}, \ldots, u^{N,o}) \leq J(u^1, u^2, \ldots, u^N), \quad \forall u \stackrel{\triangle}{=} (u^1, u^2, \ldots, u^N) \in \mathbb{U}_{rel}^{(N)}[0,T] \tag{19}$$

Any $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ satisfying (19) is called an optimal relaxed decision strategy (or control) and the corresponding $x^o(\cdot) \equiv x(\cdot; u^o(\cdot))$ (satisfying (12)) the optimal state process.
Similarly, feedback team optimal strategies are defined with respect to $u^o \in \mathbb{U}_{rel}^{(N),z}[0,T]$

(NRS): Regular Strategies. *Regular nonanticipative team optimal strategies are defined with respect to pay-off (10), constraint (9), and $u^o \in \mathbb{U}_{reg}^{(N)}[0,T]$, while feedback team optimal strategies are defined with respect to $u^o \in \mathbb{U}_{reg}^{(N),z}[0,T]$.*

By definition, Problem 1 is a dynamic team problem with each DM having a different information structure (decentralized). To the best of the authors knowledge there seems to have been no attempt in the literature to address the Problem 1. An alternative approach to handle such problems with decentralized information structures is to restrict the definition of optimality to the so-called person-by-person (player-by-player) equilibrium.
Define

$$\tilde{J}(v, u^{-i}) \stackrel{\triangle}{=} J(u^1, u^2, \ldots, u^{i-1}, v, u^{i+1}, \ldots, u^N)$$

Problem 2. *(Person-by-Person Optimality)*

(RS): Relaxed Strategies. *Given the pay-off functional (14), constraint (12) the N tuple of relaxed strategies $u^o \triangleq (u^{1,o}, u^{2,o}, \ldots, u^{N,o}) \in \mathbb{U}_{rel}^{(N)}[0,T]$ is called nonanticipative person-by-person optimal if it satisfies*

$$\tilde{J}(u^{i,o}, u^{-i,o}) = J(u^o) \leq \tilde{J}(u^i, u^{-i,o}), \quad \forall u^i \in \mathbb{U}_{reg}^i[0,T], \;\; \forall i \in \mathbb{Z}_N. \tag{20}$$

Similarly, feedback person-by-person optimal strategies are defined with respect to $u^o \in \mathbb{U}_{rel}^{(N),z}[0,T]$.

(NRS): Regular Strategies. *Regular nonanticipative person-by-person optimal strategies are defined with respect to pay-off (10), constraint (9), and $u^o \in \mathbb{U}_{reg}^{(N)}[0,T]$, while feedback person-by-person optimal strategies are defined with respect to $u^o \in \mathbb{U}_{reg}^{(N),z}[0,T]$.*

The interpretation of (20) is that the variation and hence evaluation (of team optimality) is done by the central layer and it is this layer alone that can determine if the decision for the i-th player is optimal or not. Even for Problem 2 the authors of this paper are not aware of any publication which addresses necessary and/or sufficient conditions of optimality. Conditions (20) are analogous to the Nash equilibrium strategies of team games consisting of a single pay-off and N DM. The person-by-person optimal strategy states that none of the N members (possibly with different information structures) can deviate unilaterally from the optimal strategy and gain by doing so. The rationale for the restriction to person-by-person optimal strategy is based on the fact that the actions of the N DM are not communicated to each other, and hence they cannot do better than restricting attention to this optimal strategy.

Problems 1, 2 using relaxed strategies are the main problems addressed in this paper, while conclusions for regular strategies are drawn from these results. Clearly, any strategy which is optimal for Problem 1 is also a person-by-person optimal and hence optimal for Problem 2.

III. Existence of Team Optimal Strategies

As mentioned earlier, not every control problem admits optimal regular strategies. However, in many problems relaxed strategies exist under certain mild assumptions. In this section we use

a similar procedure as the one developed in [36] for centralized information structures to prove (i) existence of solution of the distributed stochastic dynamical decision system (12), and (ii) existence of optimal relaxed strategies for the Problem 1.

A generalized sequence $u^{i,\alpha} \in \mathbb{U}^i_{rel}[0,T]$ is said to converge (in the weak* topology or) vaguely to $u^{i,o}$, written $u^{i,\alpha} \xrightarrow{v} u^{i,o}$, if and only if for every $\varphi \in L^1_{\mathcal{G}^i_T}([0,T], C(\mathbb{A}^i))$

$$\mathbb{E}\int_{[0,T]\times\mathbb{A}^i} \varphi_t(\xi) u^{i,\alpha}_t(d\xi)dt \longrightarrow \mathbb{E}\int_{[0,T]\times\mathbb{A}^i} \varphi_t(\xi) u^{i,o}_t(d\xi)dt \quad \text{as} \ \ \alpha \rightarrow \infty, \quad \forall i \in \mathbb{Z}_N.$$

With respect to the vague (weak*) topology the set $\mathbb{U}^i_{rel}[0,T]$ is compact, and from here on we assume that $\mathbb{U}^i_{rel}[0,T], \forall i \in \mathbb{Z}_N$ has been endowed with this vague topology.

Let $B^\infty_{\mathbb{F}_T}([0,T], L^2(\Omega, \mathbb{R}^n))$ denote the space of $\mathbb{F}_T$-adapted $\mathbb{R}^n$ valued second order random processes endowed with the norm topology $\parallel \cdot \parallel$ defined by

$$\parallel x \parallel^2 \stackrel{\triangle}{=} \sup_{t\in[0,T]} \mathbb{E}|x(t)|^2_{\mathbb{R}^n}.$$

To study the question of existence of solution to (12) we use the following assumptions.

Assumptions 1. *The drift f and diffusion coefficients σ associated with (12) are defined by the Borel measurable maps:*

$$f : [0,T] \times \mathbb{R}^n \times \mathbb{A}^{(N)} \longrightarrow \mathbb{R}^n, \quad \sigma : [0,T] \times \mathbb{R}^n \times \mathbb{A}^{(N)} \longrightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$$

and they are continuous in the last two arguments and assumed to satisfy the following basic properties:.

(A0) $(\mathbb{A}^i, d), \forall i \in \mathbb{Z}_N$ *are compact.*

There exists a $K \in L^{2,+}([0,T], \mathbb{R})$ *such that*

(A1) $|f(t,x,\xi) - f(t,y,\xi)|_{\mathbb{R}^n} \leq K(t)|x-y|_{\mathbb{R}^n}$ *uniformly in* $\xi \in \mathbb{A}^{(N)}$*;*

(A2) $|f(t,x,\xi)|_{\mathbb{R}^n} \leq K(t)(1+|x|_{\mathbb{R}^n})$ *uniformly in* $\xi \in \mathbb{A}^{(N)}$

(A3) $|\sigma(t,x,\xi) - \sigma(t,y,\xi)|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leq K(t)|x-y|_{\mathbb{R}^n}$ *uniformly in* $\xi \in \mathbb{A}^{(N)}$*;*

(A4) $|\sigma(t,x,\xi)|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leq K(t)(1+|x|_{\mathbb{R}^n})$ *uniformly in* $\xi \in \mathbb{A}^{(N)}$*;*

(A5) $f(t,x,\cdot), \sigma(t,x,\cdot)$ *are continuous in* $\xi \in \mathbb{A}^{(N)}$, $\forall (t,x) \in [0,T] \times \mathbb{R}^n$.

Assumptions 1, **(A1)-(A4)** are the so-called Itô conditions for existence and uniqueness of strong solutions (having continuous sample paths) [8].

The following lemma proves the existence of solutions and their continuous dependence on the decision variables.

Lemma 1. *Suppose Assumptions 1 hold. Then for any $\mathbb{F}_{0,0}$-measurable initial state x_0 having finite second moment, and any $u \in \mathbb{U}_{rel}^{(N)}[0,T]$, the following hold.*

(1) *System (12) has a unique solution $x \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega,\mathbb{R}^n))$ having a continuous modification, that is, $x \in C([0,T],\mathbb{R}^n)$, $\mathbb{P}-$a.s, $\forall i \in \mathbb{Z}_N$.*

(2) *The solution of system (12) is continuously dependent on the control, in the sense that, as $u^{i,\alpha} \xrightarrow{v} u^{i,o}$ in $\mathbb{U}_{rel}^{i}[0,T]$, $\forall i \in \mathbb{Z}_N$, $x^{\alpha} \xrightarrow{s} x^{o}$ in $B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega,\mathbb{R}^n)), \forall i \in \mathbb{Z}_N$.*

These statements also hold for feedback strategies $u \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$.

Proof: Since the class of policies $\mathbb{U}_{rel}^{i}[0,T]$, $\forall i \in \mathbb{Z}_N$ is compact in the vague topology, then $\times_{i=1}^{N}\mathbb{U}_{rel}^{i}[0,T]$ is also compact in this topology. Utilizing this observation the proof is identical to that of [36], Lemma 3.1.

■

Using the results of Lemma 1 in the next theorem we establish existence of a minimizer $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ for Problem 1. We need the following assumptions.

Assumptions 2. *The functions ℓ and φ associated with the pay-off (14) are Borel measurable maps:*

$$\ell : [0,T] \times \mathbb{R}^n \times \mathbb{A}^{(N)} \longrightarrow (-\infty,+\infty], \qquad \varphi : \mathbb{R}^n \longrightarrow (-\infty,+\infty].$$

satisfying the following basic conditions:

(B1) *$x \longrightarrow \ell(t,x,\xi)$ is continuous on $\mathbb{R}^n$ for each $t \in [0,T]$, uniformly with respect to $\xi \in \mathbb{A}^{(N)}$;*

(B2) *$\exists\ h \in L_1^{+}([0,T],\mathbb{R})$ such that for each $t \in [0,T]$, $|\ell(t,x,\xi)| \leq h(t)(1+|x|_{\mathbb{R}^n}^2)$;*

(B3) *$x \longrightarrow \varphi(x)$ is lower semicontinuous on $\mathbb{R}^n$ and $\exists\ c_0, c_1 \geq 0$ such that $|\varphi(x)| \leq c_0 + c_1|x|_{\mathbb{R}^n}^2$.*

Now we present the following existence theorem [36].

Theorem 1. *(Existence of Team Optimal Strategies) Consider Problem 1 and suppose Assumptions 1 and 2 hold. Then there exists a team decision* $u^o \stackrel{\triangle}{=} (u^{1,o}, u^{2,o}, \ldots, u^{N,o}) \in \mathbb{U}_{rel}^{(N)}[0,T]$ *at which* $J(u^1, u^2, \ldots, u^N)$ *attains its infimum. Existence also holds for* $u^o \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$.

Proof: Since the class of control policies $\mathbb{U}_{rel}^N[0,T]$ is compact in the vague topology, it suffices to prove that $J(\cdot)$ is lower semicontinuous with respect to this topology. This follows precisely from the same procedure as in [36], Theorem 3.2.

■

We conclude this section by stating that existence of team optimal strategies utilizing decentralized information structures follows directly from analogous results of centralized stochastic control strategies [13].

IV. Optimality Conditions for Relaxed Strategies

In this section we present the necessary and sufficient conditions of optimality for the team game of Problem 1. The derivation of stochastic minimum principle (necessary conditions of optimality) or stochastic Pontryagin's minimum principle is based on the martingale representation approach. For this reason we shall fisrt state certain fundamental properties of semi martingales, which are used in the derivation.

Definition 1. *Let* $\mathbb{F}_T$ *denote a complete filtration generated by an* $\mathbb{R}^m-$*dimensional Brownian motion process* $\{W(t) : t \in [0,T]\}$. *An* $\mathbb{R}^n-$*valued random process* $\{m(t) : t \in [0,T]\}$ *is said to be a square integrable continuous* $\mathbb{F}_T-$*semi martingale if and only if it has a representation*

$$m(t) = m(0) + \int_0^t v(s)ds + \int_0^t \Sigma(s)dW(s), \qquad t \in [0,T], \tag{21}$$

for some $v \in L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n)$ *and* $\Sigma \in L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ *and for some* $\mathbb{R}^n-$*valued* $\mathbb{F}_{0,0}-$*measurable random variable* $m(0)$ *having finite second moment. The set of all such semi martingales is denoted by* $\mathcal{SM}^2[0,T]$.

We need the following class of $\mathbb{F}_T-$semi martingales:

$$\mathcal{SM}_0^2[0,T] \stackrel{\triangle}{=} \Big\{ m \in \mathcal{SM}^2[0,T] : m(t) = \int_0^t v(s)ds + \int_0^t \Sigma(s)dW(s), \qquad t \in [0,T], \\ \text{for} \;\; v \in L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n) \quad \text{and} \;\; \Sigma \in L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)) \Big\}. \tag{22}$$

Now we present a fundamental result which is used in the derivation of minimum principle.

Theorem 2. *(Semi martingale Representation) The class of semi martingales $\mathcal{SM}_0^2[0,T]$ is a real linear vector space and it is a Hilbert space with respect to the norm topology $\parallel m \parallel_{\mathcal{SM}_0^2[0,T]}$ given by*

$$\parallel m \parallel_{\mathcal{SM}_0^2[0,T]} \stackrel{\triangle}{=} \Big(\mathbb{E}\int_{[0,T]} |v(t)|^2_{\mathbb{R}^n} dt + \mathbb{E}\int_{[0,T]} tr(\Sigma^*(t)\Sigma(t))dt\Big)^{1/2}.$$

Moreover, the space $\mathcal{SM}_0^2[0,T]$ is isometrically isomorphic to the space

$$L^2_{\mathbb{F}_T}([0,T],\mathbb{R}^n) \times L^2_{\mathbb{F}_T}([0,T],\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)).$$

Proof: For proof see Theorem 4.3 in [36].

■

For the derivation of stochastic minimum principle of optimality we shall require stronger regularity conditions for the drift and diffusion coefficients $\{b,\sigma\}$, as well as, for the running and terminal pay-offs functions $\{\ell,\varphi\}$. These are given below.

Assumptions 3. $\mathbb{E}|x(0)|^2_{\mathbb{R}^n} < \infty$ *and the maps of* $\{f,\sigma,\ell,\varphi\}$ *satisfy the following conditions.*

(C1) *The triple $\{f,\sigma,\ell\}$ are measurable in $t \in [0,T]$;*

(C2) *The quadruple $\{f,\sigma,\ell,\varphi\}$ are once continuously differentiable with respect to the state variable $x \in \mathbb{R}^n$;*

(C3) *The first derivatives of $\{f,\sigma\}$ with respect to the state are bounded uniformly on $[0,T] \times \mathbb{R}^n \times \mathbb{A}^{(N)}$.*

Consider the Gateaux derivative of σ with respect to the variable at the point $(t,z,\nu) \in [0,T] \times \mathbb{R}^n \times_{i=1}^N \mathcal{M}_1(\mathbb{A}^i)$ in the direction $\eta \in \mathbb{R}^n$ defined by

$$\sigma_x(t,z,\nu;\eta) \stackrel{\triangle}{=} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\Big\{\sigma(t,z+\varepsilon\eta,\nu) - \sigma(t,z,\nu)\Big\}, \quad t \in [0,T].$$

Note that the map $\eta \longrightarrow \sigma_x(t,z,\nu;\eta)$ is linear, and it follows from Assumptions 3, **(C3)** that there exists a finite positive number $\beta > 0$ such that

$$|\sigma_x(t,z,\nu;\eta)|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)} \leq \beta|\eta|_{\mathbb{R}^n}, \quad t \in [0,T].$$

In order to present the necessary conditions of optimality we need the so called variational equation. Let us first introduce the variational equation for nonanticipative information structures. Suppose $u^o \stackrel{\triangle}{=} (u^{1,o},u^{2,o},\dots,u^{N,o}) \in \mathbb{U}_{rel}^{(N)}[0,T]$ denotes the optimal decision and $u \stackrel{\triangle}{=}$

$(u^1, u^2, \ldots, u^N) \in \mathbb{U}_{rel}^{(N)}[0,T]$ any other decision. Since $\mathbb{U}_{rel}^i[0,T]$ is convex $\forall i \in \mathbb{Z}_N$, it is clear that for any $\varepsilon \in [0,1]$,

$$u_t^{i,\varepsilon} \stackrel{\triangle}{=} u_t^{i,o} + \varepsilon(u_t^i - u_t^{i,o}) \in \mathbb{U}_{rel}^i[0,T], \quad \forall i \in \mathbb{Z}_N.$$

Let $x^\varepsilon(\cdot) \equiv x^\varepsilon(\cdot; u^\varepsilon(\cdot))$ and $x^o(\cdot) \equiv x^o(\cdot; u^o(\cdot)) \in B_{\mathbb{F}_T}^\infty([0,T], L^2(\Omega, \mathbb{R}^n))$ denote the solutions of the system equation (12) corresponding to $u^\varepsilon(\cdot)$ and $u^o(\cdot)$, respectively. Consider the limit

$$Z(t) \stackrel{\triangle}{=} \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\Big\{x^\varepsilon(t) - x^o(t)\Big\}, \quad t \in [0,T].$$

We have the following result characterizing the process $\{Z(t) : t \in [0,T]\}$.

Lemma 2. *Suppose Assumptions 3 hold and consider nonanticipative strategies $\mathbb{U}_{rel}^{(N)}[0,T]$. The process $\{Z(t) : t \in [0,T]\}$ as defined above is an element of the Banach space $B_{\mathbb{F}_T}^\infty([0,T], L^2(\Omega, \mathbb{R}^n))$ and it is the unique solution of the variational stochastic differential equation*

$$\begin{aligned} dZ(t) &= f_x(t, x^o(t), u_t^o)Z(t)dt + \sigma_x(t, x^o(t), u_t^o; Z(t))\ dW(t) \\ &+ \sum_{i=1}^N f(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o})dt + \sum_{i=1}^N \sigma(t, x^o(t), u_t^{-i,o}, u_t^i - u_t^{i,o})dW(t), \quad Z(0) = 0. \end{aligned} \tag{23}$$

having a continuous modification.

Proof: We closely follow the steps in [33]. Writing the system (12) as an integral equation with solutions x^ε, x^o corresponding to controls u^ε, u^o respectively and taking the difference $x^\varepsilon(t) - x^o(t)$ and dividing by ε and then letting $\varepsilon \longrightarrow 0$, it can be shown that it converges for all $t \in [0,T], \mathbb{P}-a.s.$ to the solution of system (23). Note that the system (23) is a linear stochastic differential equation in Z with non homogeneous terms given by the sum of the last two terms. Let $\{z(t) : t \in [0,T]\}$ denote the solution of its homogenous part given by

$$dz(t) = f_x(t, x^o(t), u_t^o)z(t)dt + \sigma_x(t, x^o(t), u_t^{i,o}; z(t))dW(t), \quad z(s) = \zeta, \quad t \in [s,T]. \tag{24}$$

By Assumptions 3 and Lemma 1 this system has a unique solution $\{z(t) : t \in [s,T]\}$ given by

$$z(t) = \Psi(t,s)\zeta, \quad t \in [s,T],$$

where $\Psi(t,s), t \in [s,T]$ is the random ($\mathbb{F}_T-$adapted) transition operator for the homogenous system. Since the derivatives of f and σ with respect to the state are uniformly bounded, the

transition operator $\Psi(t,s), t \in [s,T]$ is uniformly $\mathbb{P}-$a.s. bounded (with values in the space of $n \times n$ matrices).

By Using the random transition operator Ψ we can write the solution of the non homogenous stochastic differential equation (23) as follows,

$$Z(t) = \int_0^t \Psi(t,s) d\eta(s), \quad t \in [0,T], \tag{25}$$

where $\{\eta(t) : t \in [0,T]\}$ is the semi martingale given by the following Ito differential,

$$\begin{aligned} d\eta(t) = &\sum_{i=1}^N f(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) dt \\ &+ \sum_{i=1}^N \sigma(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o})\ dW(t), \quad \eta(0) = 0, \quad t \in (0,T]. \end{aligned} \tag{26}$$

Note that $\{\eta(t) : t \in [0,T]\}$ is a continuous square integrable $\mathbb{F}_T-$adapted semi martingale. The fact that it has continuous modification follows directly from the representation (25) and the continuity of the semi martingale $\{\eta(t) : t \in [0,T]\}$. ∎

Clearly, the variational equation for nonanticipative strategies $\mathbb{U}_{rel}^{(N)}[0,T]$ is obtained as in centralized control strategies found in [36]. Next, we discuss the variational equation for feedback information structures. For $u \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$ the variational equation will also involve derivatives of u with respect to the state trajectory x, since such strategies utilize feedback. To avoid this technicality, we first address the question as to whether optimizing $J(u)$ over nonanticipative information structures is the same as optimizing $J(u)$ over feedback information structures. If this is the case then the variational equation for $u \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$ will be that of $u \in \mathbb{U}_{rel}^{(N)}[0,T]$. We shall require the following assumption.

Assumptions 4. *The following holds.*

(E1) *The diffusion coefficient σ is independent of u and both $\sigma(\cdot,\cdot)$ and $\sigma^{-1}(\cdot,\cdot)$ are uniformly bounded.*

Under the (additional) Assumptions 4 we can prove the following theorem.

Theorem 3. *Consider Problem 1 and suppose Assumptions 1 and 4 hold. Define the $\sigma-$algebras*

$$\mathcal{F}_{0,t}^{x(0),W} \triangleq \sigma\{x(0), W(s) : 0 \leq s \leq t\}, \quad \mathcal{F}_{0,t}^{x^u} \triangleq \sigma\{x^u(s) : 0 \leq s \leq t\}, \quad \forall t \in [0,T].$$

Then for all $u \in \mathbb{U}_{rel}^{(N),x^u}[0,T]$ *the two* σ*-algebras are equivalent written as an equality,* $\mathcal{F}_{0,t}^{x(0),W} = \mathcal{F}_{0,t}^{x^u}, \forall t \in [0,T]$.

Proof: Clearly, by Lemma 1, we have $\mathcal{F}_{0,t}^{x^u} \subset \mathcal{F}_{0,t}^{x(0),W}, \forall u \in \mathbb{U}_{rel}^{(N)}[0,T], t \in [0,T]$. By use of Assumptions 4 one can easily verify that $\mathcal{F}_{0,t}^{x(0),W} \subset \mathcal{F}_{0,t}^{x^u}, \forall t \in [0,T]$. This completes the proof. ■

Under the conditions of Theorem 3, for any stochastic kernel $\{u_t^i(\Gamma) \equiv q_t^i(\Gamma|\mathcal{G}_{0,t}^{x^{i,u}}) : t \in [0,T]\} \in \mathbb{U}_{rel}^{x^{i,u}}[0,T], \Gamma \in \mathcal{B}(\mathbb{A}^i)$ which is $\mathcal{G}_{0,t}^{x^{i,u}}-$measurable there exists a function $\phi^i(\cdot)$ adapted to a sub-$\sigma-$algebra of $\mathcal{F}_{0,t}^i \subset \mathcal{F}_{0,t}^{x(0),W}$ such that $u_t^i(\Gamma) = q_t^i(\Gamma|\phi^i(t, x(0), W(\cdot \bigwedge t, \omega))), \mathbb{P}-a.s, \forall t \in [0,T], i = 1, \ldots N$.
Let $\mathcal{F}_T^i \stackrel{\triangle}{=} \{\mathcal{F}_{0,t}^i : t \in [0,T]\}, \mathcal{G}_T^{x^{i,u}} \stackrel{\triangle}{=} \{\mathcal{G}_{0,t}^{x^{i,u}} : t \in [0,T]\}, i = 1, \ldots, \mathbb{Z}_N$, and define all such adapted nonanticipative functions by

$$\overline{\mathbb{U}}_{rel}^i[0,T] \stackrel{\triangle}{=} \Big\{u \in L_{\mathcal{F}_T^i}^\infty([0,T], \mathcal{M}_1(\mathbb{A}^i)) : \ u^i \in L_{\mathcal{G}_T^{x^{i,u}}}^\infty([0,T], \mathcal{M}_1(\mathbb{A}^i))\Big\}, \ \forall i \in \mathbb{Z}_N. \tag{27}$$

Next, we introduce the following additional assumptions.

Assumptions 5. *The following holds.*

(E2) $\mathbb{U}_{rel}^{x^{i,u}}[0,T]$ *is dense in* $\overline{\mathbb{U}}_{rel}^i[0,T], \forall i \in \mathbb{Z}_N$.

Under the additional Assumptions 5 we can prove the following result.

Theorem 4. *Consider Problem 1 with control strategies from* $\mathbb{U}_{rel}^{(N),x^u}[0,T]$. *Under Assumptions 1, 2, 5, and* $|\varphi_x(x) + \ell_x(t,x,u)|_{\mathbb{R}^n} \leq K(1 + |x|_{\mathbb{R}^n})$ *we have,*

$$\inf_{u \in \times_{i=1}^N \overline{\mathbb{U}}_{rel}^i[0,T]} J(u) = \inf_{u \in \times_{i=1}^N \mathbb{U}_{rel}^{z^{i,u}}[0,T]} J(u). \tag{28}$$

Proof: The assertion is obvious because of the density assumption (E2) and the continuity of J in the vague topology.

■

The point to be made regarding Theorem 4 is that if $u \in \mathbb{U}_{rel}^{(N),x^u}[0,T]$ achieves the infimum of $J(u)$ then it is also optimal with respect to $\overline{\mathbb{U}}_{rel}^{(N)}[0,T] \stackrel{\triangle}{=} \times_{i=1}^N \overline{\mathbb{U}}_{rel}^i[0,T]$. Consequently, the necessary conditions for feedback information structures $u \in \mathbb{U}_{rel}^{(N),x^u}[0,T]$ to be optimal are those for which nonanticipative information structures $u \in \overline{\mathbb{U}}_{rel}^{(N)}[0,T]$ are optimal.

In the next remark we give an example for which Assumptions 5 hold, and hence Theorem 4 is valid.

Remark 1. *Suppose x^1 and x^2 are governed by the following stochastic differential equations*

$$dx^1(t) = f^1(t,x^1(t),u^1(t))dt + \sigma^1(t,x^1(t))dW^1(t), \quad x^1(0) = x^1_0, \tag{29}$$

$$dx^2(t) = f^2(t,x^1(t),x^2(t),u^1(t),u^2(t))dt + \sigma^2(t,x^1(t),x^2(t))dW^2(t), \quad x^2(0) = x^2_0, \tag{30}$$

$$z^1(t) = h^1(t,x^1(t)), \quad z^2(t) = h^2(t,x^1(t),x^2(t)), \quad t \in [0,T], \tag{31}$$

where h^1, h^2 are measurable, $W^1(\cdot), W^2(\cdot)$ are independent, and $u^1 \in \mathbb{U}_{rel}^{1,z^{1,u^1}}[0,T], u^2 \in \mathbb{U}_{rel}^{2,z^{2,u^2}}[0,T]$. If we further assume that $\{\sigma^i(\cdot,\cdot)\}$ and their inverses are bounded, then we can find $\overline{\mathbb{U}}^i_{rel}[0,T], i=1,2$ for which (E2) holds, and thus Theorem 4 holds. The structure of the stochastic dynamics (29), (30) can be generalized to more than two coupled systems.

Next, we introduce the following alternative theorem to Theorem 4, which does not employ Assumptions 5.

Theorem 5. *Consider Problem 1 with strategies from $\mathbb{U}_{rel}^{(N),z^u}[0,T]$, under Assumptions 1, 2, III, 4 and $|\varphi_x(x) + \ell_x(t,x,u)|_{\mathbb{R}^n} \leq K(1+|x|_{\mathbb{R}^n})$.*
Then $\mathbb{U}_{rel}^{z^{i,u}}[0,T]$ is dense in $\overline{\mathbb{U}}^i_{rel}[0,T], \forall i \in \mathbb{Z}_N$ and

$$\inf_{u \in \times_{i=1}^N \overline{\mathbb{U}}^i_{rel}[0,T]} J(u) = \inf_{u \in \times_{i=1}^N \mathbb{U}_{rel}^{z^{i,u}}[0,T]} J(u). \tag{32}$$

Proof: The derivation is based on [40] but extended to relaxed strategies. By Theorem 3, for any $u^i \in \mathbb{U}_{rel}^{z^{i,u}}[0,T]$ which is $\mathcal{G}_T^{z^{i,u}}-$adapted we can define the set $\overline{\mathbb{U}}^i_{rel}[0,T], i=1,\ldots,N$ via (27). For any $u \in \overline{\mathbb{U}}^{(N)}_{rel}[0,T] \triangleq \times_{i=1}^N \overline{\mathbb{U}}^i_{rel}[0,T]$, $k = \frac{T}{M}$, and any test function $\phi \in C(\mathbb{A}^{(N)})$, define

$$u_{k,t}[\phi] \triangleq \begin{cases} \int_{\mathbb{A}^{(N)}} \phi(\xi) u_0(d\xi) & \text{for} \quad 0 \leq t < k \qquad u_0 \in \mathbb{A}^{(N)} \\ \frac{1}{k}\int_{(n-1)k}^{nk} \int_{\mathbb{A}^{(N)}} \phi(\xi) u_s(d\xi) ds & \text{for} \quad nk \leq t < (n+1)k, \quad n = 1,\ldots,M-1. \end{cases} \tag{33}$$

Clearly $u_k \in \overline{\mathbb{U}}^{(N)}_{rel}[0,T]$, and $u_k \longrightarrow u$ in $L^\infty_{\mathcal{F}_T}([0,T], \mathcal{M}_1(\mathbb{A}^{(N)}))$ in the weak star sense. We need to show that $u_k \in \mathbb{U}^{(N),z^{u_k}}[0,T]$. Let x_k denote the trajectory corresponding to u_k, and $\mathcal{F}_{0,t}^{x_k^u}$ the $\sigma-$algebra generated by $\{x_k(s) : 0 \leq s \leq t\}$. Define

$$I_k(t) \triangleq \int_0^t \sigma(s,x_k(s))dW(t) = x_k(t) - x(0) - \int_0^t f(s,x_k(s),u_k(s))ds, \tag{34}$$

and

$$W(t)=\int_0^t \sigma(s,x_k(s))^{-1}dI_k(s). \tag{35}$$

Since $u_k \in \overline{\mathbb{U}}_{rel}^{(N)}[0,T]$, the process $I_k(t)$ is $\mathcal{F}_{0,t}^{x_k^u}-$measurable, for $0 \leq t < k$. Hence,

$$\mathcal{F}_{0,t}^{x(0),W}=\mathcal{F}_{0,t}^{x_k^u}, \quad 0 \leq t \leq k. \tag{36}$$

Therefore, $u_{k,t}$ is $\mathcal{F}_{0,t}^{x_k^u}-$ measurable for $k \leq t \leq 2k$. From the above equations it follows that (36) also holds for $k \leq t \leq 2k$, and by induction that $\mathcal{F}_{0,t}^{x(0),W} = \mathcal{F}_{0,t}^{x_k^u}, \forall t \in [0,T]$. Therefore, $u_{k,t}^i$ is also (weak star) measurable with respect to $\mathcal{F}_{0,t}^{x_k^u}$. Hence , for any u_t^i which is (weak star) measurable with respect to a nonanticipative functional $z^i = h^i(t,x)$ there exists a nonanticipative functional of $\{x(0),W\}$ which realizes it. By Theorem 4 the derivation is complete. ∎

.

Before we prove the optimality conditions we define the Hamiltonian system of equations. The Hamiltonian is a real valued function

$$\mathbb{H}:[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)\times\mathcal{M}_1(\mathbb{A}^{(N)})\longrightarrow\mathbb{R}$$

given by

$$\mathbb{H}(t,\xi,\zeta,M,\nu)\stackrel{\triangle}{=}\langle f(t,\xi,\nu),\zeta\rangle + tr(M^*\sigma(t,\xi,\nu))+\ell(t,\xi,\nu), \quad t\in[0,T]. \tag{37}$$

For any $u \in \mathbb{U}_{rel}^{(N)}[0,T]$, the adjoint process is $(\psi,Q)\in L^2_{\mathbb{F}_T}([0,T],\mathbb{R}^n)\times L^2_{\mathbb{F}_T}([0,T],\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n))$ satisfies the following backward stochastic differential equation

$$\begin{aligned} d\psi(t) &= -f_x^*(t,x(t),u_t)\psi(t)dt - V_Q(t)dt - \ell_x(t,x(t),u_t)dt + Q(t)dW(t), \quad t\in[0,T), \\ &= -\mathbb{H}_x(t,x(t),\psi(t),Q(t),u_t)dt + Q(t)dW(t), \end{aligned} \tag{38}$$

$$\psi(T)=\varphi_x(x(T)) \tag{39}$$

where $V_Q \in L^2_{\mathbb{F}_T}([0,T],\mathbb{R}^n)$ is given by $\langle V_Q(t),\zeta\rangle = tr(Q^*(t)\sigma_x(t,x(t),u_t;\zeta)), t\in[0,T]$ (e.g., $V_Q(t)=\sum_{k=1}^m\left(\sigma_x^{(k)}(t,x(t),u_t)\right)^*Q^{(k)}(t), \quad t\in[0,T]$, $\sigma^{(k)}$ is the kth column of σ, $\sigma_x^{(k)}$ is the derivative of $\sigma^{(k)}$ with respect to the state, for $k=1,2,\ldots,m$, $Q^{(k)}$ is the kth column of Q). In terms of the Hamiltonian, the state process satisfies the stochastic differential equation

$$\begin{aligned} dx(t) &= f(t,x(t),u_t)dt + \sigma(t,x(t),u_t)dW(t), \quad t\in(0,T], \\ &= \mathbb{H}_\psi(t,x(t),\psi(t),Q(t),u_t)dt + \sigma(t,x(t),u_t)dW(t), \end{aligned} \tag{40}$$

$$x(0)=x_0 \tag{41}$$

A. Necessary Conditions of Optimality

In this section we state and prove the necessary conditions for team optimality. Specifically, given that $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ or $u^o \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$ is team optimal, we show that it leads naturally to the Hamiltonian system of equations (called necessary conditions). The derivation is based on the semi martingale representation as in [36] with some modifications necessary to admit decentralized strategies adapted to an arbitrary filtration.

In the following theorem we present the necessary conditions of optimality for Problem 1.

Theorem 6. *(Necessary conditions for team optimality) Consider Problem 1 under Assumptions 2, 3.*

(I) *Suppose* $\mathbb{F}_T = \sigma\{x(0), W(t), t \in [0,T]\}$ *and* $\mathbb{U}_{rel}^{(N)}[0,T]$ *is the class of relaxed controls adapted to this filtration. For an element* $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ *with the corresponding solution* $x^o \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ *to be team optimal, it is necessary that the following conditions hold.*

  (1) *There exists a semi martingale* $m^o \in \mathcal{SM}_0^2[0,T]$ *with the intensity process* $(\psi^o, Q^o) \in L_{\mathbb{F}_T}^2([0,T], \mathbb{R}^n) \times L_{\mathbb{F}_T}^2([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$.

  (2) *The processes* $\{u^o, x^o, \psi^o, Q^o\}$ *satisfy the inequality :*

$$\sum_{i=1}^N \mathbb{E} \int_0^T \mathbb{H}(t, x^o(t)\psi^o(t), Q^o(t), u_t^{-i,o}, u_t^i - u_t^{i,o}) dt \geq 0, \quad \forall u \in \mathbb{U}_{rel}^{(N)}[0,T]. \tag{42}$$

  (3) *The process* (ψ^o, Q^o) *is the unique solution of the backward stochastic differential equation (38), (39) and that, for* $\mathcal{G}_{0,t}^i \subset \mathbb{F}_{0,t}$*, the control* $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ *satisfies the point wise almost sure inequalities.*

$$\mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^0(t), Q^o(t), u_t^{-i,o}, \nu^i) | \mathcal{G}_{0,t}^i\Big\} \geq \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^o) | \mathcal{G}_{0,t}^i\Big\},$$
$$\forall \nu^i \in \mathcal{M}_1(\mathbb{A}^i), a.e. t \in [0,T], \mathbb{P}|_{\mathcal{G}_{0,t}^i} - a.s., i = 1,2,\ldots,N. \tag{43}$$

(II) *Suppose* $\mathbb{F}_T$ *is as above, and the Assumption 5 holds. For an element* $u^o \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$ *with the corresponding solution* $x^o \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ *to be team optimal, it is necessary that the statements of Part* **(I)** *hold with* $\mathcal{G}_{0,t}^i$ *replaced by* $\mathcal{G}_{0,t}^{z^{i,u}}, \forall t \in [0,T]$.

Proof: The derivation of **(1), (2)** follows closely the basic steps of centralized strategies in [36], from which the derivation of team necessary conditions of optimality **(3)** are established. **(I)**. **(1)** Suppose $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ is an optimal team decision and $u \in \mathbb{U}_{rel}^{(N)}[0,T]$ any other admissible decision. Since $\mathbb{U}_{rel}^{i}[0,T]$ is convex $\forall i \in \mathbb{Z}_N$, we have, for any $\varepsilon \in [0,1]$, $u_t^{i,\varepsilon} \triangleq u_t^{i,o} + \varepsilon(u_t^i - u_t^{i,o}) \in \mathbb{U}_{rel}^{i}[0,T], \forall i \in \mathbb{Z}_N$. Let $x^\varepsilon(\cdot) \equiv x^\varepsilon(\cdot; u^\varepsilon(\cdot)), x^o(\cdot) \equiv x^o(\cdot; u^o(\cdot)) \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ denote the solutions of the system (12) and (41) corresponding to $u^\varepsilon(\cdot)$ and $u^o(\cdot)$, respectively. Since $u^o(\cdot) \in \mathbb{U}_{rel}^{(N)}[0,T]$ is optimal it is clear that

$$J(u^\varepsilon) - J(u^o) \geq 0, \quad \forall \varepsilon \in [0,1], \;\; \forall u \in \mathbb{U}_{rel}^{(N)}[0,T]. \tag{44}$$

Define the Gateaux differential of J at u^o in the direction $u - u^o$ by

$$dJ(u^o, u - u^0) \triangleq \lim_{\varepsilon \downarrow 0} \frac{J(u^\varepsilon) - J(u^o)}{\varepsilon} \equiv \frac{d}{d\varepsilon} J(u^\varepsilon)|_{\varepsilon=0}.$$

Dividing the expression (44) by ε and letting $\varepsilon \downarrow 0$ we obtain

$$dJ(u^o, u - u^0) = L(Z) + \sum_{i=1}^N \mathbb{E} \int_0^T \ell(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) dt \geq 0, \quad \forall u \in \mathbb{U}_{rel}^{(N)}[0,T], \tag{45}$$

where $L(Z)$ is given by the functional

$$L(Z) = \mathbb{E}\Big\{ \int_0^T \langle \ell_x(t, x^o(t), u_t^o), Z(t) \rangle \; dt + \langle \varphi_x(x^o(T)), Z(T) \rangle \Big\}. \tag{46}$$

Since by Lemma 2, the process $Z(\cdot) \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ and it is also continuous $\mathbb{P}-$a.s it follows from Assumptions 2, **(B2)**, and Assumptions 3, that $Z \longrightarrow L(Z)$ is a continuous linear functional. Further, by Lemma 2, $\eta \longrightarrow Z$ is a continuous linear map from the Hilbert space $\mathcal{SM}_0^2[0,T]$ to the B-space $B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ given by the expression (25). Thus the composition map $\eta \longrightarrow Z \longrightarrow L(Z) \equiv \tilde{L}(\eta)$ is a continuous linear functional on $\mathcal{SM}_0^2[0,T]$. Then by virtue of Riesz representation theorem for Hilbert spaces, there exists a semi martingale $m^o \in \mathcal{SM}_0^2[0,T]$ with intensity $(\psi^o, Q^o) \in L_{\mathbb{F}_T}^2([0,T], \mathbb{R}^n) \times L_{\mathbb{F}_T}^2([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ such that

$$\begin{aligned} L(Z) \triangleq \tilde{L}(\eta) = (m^o, \eta)_{\mathcal{SM}_0^2[0,T]} &= \sum_{i=1}^N \mathbb{E} \int_0^T \langle \psi^o(t), f(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) \rangle dt \\ &+ \sum_{i=1}^N \mathbb{E} \int_0^T tr(Q^{o,*}(t) \sigma(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o})) dt. \end{aligned} \tag{47}$$

This proves **(1)**.

(2) Substituting (47) into (45) we obtain the following variational equation.

$$
\begin{aligned}
dJ(u^o, u-u^0) = &\sum_{i=1}^N \mathbb{E} \int_0^T \langle \psi^o(t), f(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) \rangle dt \\
&+ \sum_{i=1}^N \mathbb{E} \int_0^T tr(Q^{o,*}(t) \sigma(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o})) dt \\
&+ \sum_{i=1}^N \mathbb{E} \int_0^T \ell(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) dt \geq 0, \quad \forall u \in \mathbb{U}_{rel}^{(N)}[0,T].
\end{aligned} \tag{48}
$$

It follows from the definition of the Hamiltonian that the inequality (48) is precisely (42) along with the pair $\{(\psi^o(t), Q^o(t)) : t \in [0,T]\}$. This completes the proof of **(2)**.

(3) Next, we prove that the pair $\{(\psi^o(t), Q^o(t)) : t \in [0,T]\})$ is given by the solution of the adjoint equations (38), (39). Computing the Itô differential of the scalar product $\langle Z, \psi^o \rangle$ and integrating this over $[0,T]$, it follows from the variational equation (23) that

$$
\begin{aligned}
\mathbb{E}\langle Z(T), \psi^o(T) \rangle = &\mathbb{E}\Big\{ \int_0^T \langle Z(t), f_x^*(t, x^o(t), u_t^o) \psi^o(t) dt + \sigma_x^*(t, x^o, u_t^o; \psi^o) dW(t) + d\psi^o(t) \rangle \\
&+ \sum_{i=1}^N \int_0^T \langle f(t, x^o(t), u_t^{-i,o}, u_t^i - u_t^{i,o}), \psi^o(t) \rangle dt \\
&+ \sum_{i=1}^N \int_0^T \langle \sigma^*(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}) \psi^o(t), dW(t) \rangle + \int_0^T < dZ, d\psi^o > (t) \Big\}
\end{aligned} \tag{49}
$$

$$
\begin{aligned}
&= \mathbb{E}\Big\{ \int_0^T \langle Z(t), f_x^*(t, x^o(t), u_t^o) \psi^o(t) dt + d\psi^o(t) \rangle \\
&+ \sum_{i=1}^N \int_0^T \langle f(t, x^o(t), u^{-i,o}, u_t^i - u_t^{i,o}), \psi^o(t) \rangle dt \\
&\int_0^T < dZ, d\psi^o > (t) \Big\},
\end{aligned} \tag{50}
$$

where the last bracket $< \cdot, \cdot >$ in each of the above expressions is the quadratic variation between the two processes, and the stochastic integrals in (49) have zero expectation giving (50). Since Itô derivatives of the variation process $\{Z(t) : t \in [0,T]\}$ and the adjoint process $\{\psi^o(t) : t \in [0,T]\}$

have the form

$$dZ(t) = \text{bounded variation terms} + \sigma_x(t, x^o(t), u^o_t; Z(t))dW(t)$$
$$+ \sum_{i=1}^{N} \sigma(t, x^o(t), u_t^{-i,o}, u^i_t - u^{i,o}_t)dW(t), \quad Z(0) = 0, \quad t \in (0, T], \tag{51}$$
$$d\psi^o(t) = \text{bounded variation terms} + Q^o(t)dW(t), \quad \psi^o(T) = \varphi_x(x^o(T)), \tag{52}$$

their quadratic variation is given by

$$\mathbb{E}\int_0^T < dZ, d\psi^o > (t) = \mathbb{E}\Big\{ \int_0^T tr(Q^{o,*}\sigma_x(t, x^o(t), u^o_t; Z(t)))dt\Big\}$$
$$+ \sum_{i=1}^{N} \mathbb{E}\Big\{ \int_0^T tr(Q^{o,*}(t)\sigma(t, x^o(t), u_t^{-i,o}, u^i_t - u^{i,o}_t))dt\Big\}. \tag{53}$$

The first term on the right hand side of the above expression is linear in Z, hence there exists a process $\{V_{Q^o}(t) : t \in [0, T]\}$, given by the following expression

$$\langle V_{Q^o}(t), Z(t)\rangle \stackrel{\triangle}{=} tr(Q^{o,*}(t)\sigma_x(t, x^o(t), u^o_t; Z(t))). \tag{54}$$

By Assumptions 3, σ has uniformly bounded spatial first derivative and it follows from the semi martingale representation that $Q^o \in L^2_{\mathbb{F}_T}([0, T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ and hence $V_{Q^o} \in L^2_{\mathbb{F}_T}([0, T], \mathbb{R}^n)$. Substituting (54) into (53) and (53) into (50), we obtain

$$\mathbb{E}(Z(T), \psi^o(T)) = \mathbb{E}\Big\{\int_0^T \langle Z(t), f^*_x(t, x^o(t), u^o_t)\psi^o dt + V_{Q^o}(t)dt - Q^o(t)dW(t) + d\psi^o(t)\rangle\Big\}$$
$$+ \sum_{i=1}^{N} \mathbb{E}\Big\{\int_0^T \langle f(t, x^o(t), u_t^{-i,o}, u^i_t - u^{i,o}_t), \psi^o(t)\rangle dt$$
$$+ tr(Q^{o,*}(t)\sigma(t, x^o(t), u_t^{-i,o}, u^i_t - u^{i,o}_t))dt\Big\}. \tag{55}$$

Thus, by setting

$$d\psi^o(t) = -f^*_x(t, x^o(t), u^o_t)\psi^o(t)dt - V_{Q^o}(t)dt + Q^o(t)dW(t) - \ell_x(t, x^o(t), u^o_t)dt, \quad t \in [0, T) \tag{56}$$

$$\psi^o(T) = \varphi_x(x^o(T)), \tag{57}$$

it follows from (55) and the expression for the functional $L(\cdot)$ given by (46) that

$$\begin{aligned} L(Z) &= \mathbb{E}\Big\{ \langle Z(T), \psi^o(T)\rangle + \int_0^T \langle Z(t), \ell_x(t,x^o(t),u_t^o)\rangle dt \Big\} \\ &= \sum_{i=1}^N \mathbb{E}\Big\{ \int_0^T \langle f(t,x^o(t),u_t^{-i,o}, u_t^i - u_t^{i,o}), \psi^o(t)\rangle + tr(Q^{o,*}\sigma(t,x^o(t),u_t^{-i,o},u_t^i - u_t^{i,o}))dt \Big\}. \end{aligned} \tag{58}$$

Substituting (58) into (45) we again obtain (42), as expected. This is precisely what was obtained by the semi martingale argument giving (47). Thus the pair $\{(\psi^o(t), Q^o(t)) : t \in [0,T]\}$ must satisfy the backward stochastic differential equation (56), (57), which is precisely the adjoint equation given by (38), (39). Since ψ^o satisfies the stochastic differential equation and T is finite, it follows from the classical theory of Itô differential equations that ψ^o is actually an element of $B^\infty_{\mathbb{F}_T}([0,T], L^2(\Omega,\mathbb{R}^n)) \subset L^2_{\mathbb{F}_T}([0,T],\mathbb{R}^n)$. In other words, ψ^o is more regular than predicted by semi martingale theory. Hence, by our Assumptions on σ it is easy to verify that $\sigma_x^*(t,x^o(t),u_t^o;\psi^o(t)) \in L^2_{\mathbb{F}_T}([0,T],\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n))$ and

$$\sigma^*(t,x^o(t),u_t^{-i,o},u_t^i - u_t^{i,o})\psi^o(t) \in L^2_{\mathbb{F}_T}([0,T],\mathbb{R}^n), \quad i = 1,\ldots,N.$$

This proves the first part of **(3)**.

Now we show (43). Write (42) in terms of the Hamiltonian as follows.

$$\sum_{i=1}^N \mathbb{E}\Big\{ \int_0^T \mathbb{H}(t,x^o(t),\psi^o(t),Q^o(t),u_t^{-i,o},u_t^i - u_t^{i,o})dt \Big\} \geq 0, \quad \forall u \in \mathbb{U}_{rel}^{(N)}[0,T], \tag{59}$$

where the triple $\{x^o,\psi^o,Q^o\}$ is the unique solution of the Hamiltonian system (38), (39), (40), (41). By using the property of conditional expectation then

$$\sum_{i=1}^N \mathbb{E}\Big\{ \int_0^T \mathbb{E}\Big\{ \mathbb{H}(t,x^o(t),\psi^o(t),Q^o(t),u_t^{-i,o},u_t^i - u_t^{i,o}) | \mathcal{G}^i_{0,t}\Big\} dt \Big\} \geq 0, \ \ \forall u \in \mathbb{U}_{rel}^{(N)}[0,T]. \tag{60}$$

Let $t \in (0,T)$, $\omega \in \Omega$ and $\varepsilon > 0$, and consider the sets $I^i_\varepsilon \equiv [t,t+\varepsilon] \subset [0,T]$ and $\Omega^i_\varepsilon(\subset \Omega) \in \mathcal{G}^i_{0,t}$ containing ω such that $|I^i_\varepsilon| \to 0$ and $\mathbb{P}(\Omega^i_\varepsilon) \to 0$ as $\varepsilon \to 0$, for $i = 1,2,\ldots,N$. For any sub-sigma algebra $\mathcal{G} \subset \mathbb{F}$, let $\mathbb{P}|_{\mathcal{G}}$ denote the restriction of the probability measure $\mathbb{P}$ on to the σ-algebra $\mathcal{G}$. For any (vaguely) $\mathcal{G}^i_{0,t}-$adapted $\nu^i \in \mathcal{M}_1(\mathbb{A}^i)$, construct

$$u_t^i = \begin{cases} \nu^i & \text{for} \quad (t,\omega) \in I^i_\varepsilon \times \Omega^i_\varepsilon \\ u_t^{i,o} & \text{otherwise} \end{cases} \qquad i = 1,2,\ldots,N. \tag{61}$$

Clearly, it follows from the above construction that $u^i \in \mathbb{U}^i_{rel}[0,T]$. Substituting (61) in (60) we obtain the following inequality

$$\sum_{i=1}^N \int_{\Omega^i_\varepsilon \times I^i_\varepsilon} \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{-i,o}, \nu^i - u_t^{i,o})|\mathcal{G}^i_{0,t}\Big\}dt \geq 0,$$

$$\forall \nu^i \in \mathcal{M}_1(\mathbb{A}^i), a.e. t \in [0,T], \mathbb{P}|_{\mathcal{G}^i_{0,t}} - a.s., \quad i = 1, 2, \ldots, N. \tag{62}$$

Letting $|I^i_\varepsilon|$ denote the Lebesgue measure of the set I^i_ε and dividing the above expression by the product measure $\mathbb{P}(\Omega^i_\varepsilon)|I^i_\varepsilon|$ and letting $\varepsilon \to 0$ we arrive at the following inequality.

$$\sum_{i=1}^N \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{-i,o}, \nu^i)|\mathcal{G}^i_{0,t}\Big\} \geq \sum_{i=1}^N \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{-i,o}, u_t^{i,o})|\mathcal{G}^i_{0,t}\Big\},$$

$$\forall \nu^i \in \mathcal{M}_1(\mathbb{A}^i), a.e. t \in [0,T], \mathbb{P}|_{\mathcal{G}^i_{0,t}} - a.s., i = 1, 2, \ldots, N. \tag{63}$$

To complete the proof of **(3)** define

$$g^i(t,\omega) \triangleq \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{-i,o}, \nu^i - u_t^{i,o})|\mathcal{G}^i_{0,t}\Big\}, \quad t \in [0,T], \quad \forall i \in \mathbb{Z}_N. \tag{64}$$

We shall show that

$$g^i(t,\omega) \geq 0, \quad \forall \nu^i \in \mathcal{M}_1(\mathbb{A}^i), \; a.e. \; t \in [0,T], \quad \mathbb{P}|_{\mathcal{G}^i_{0,t}} - a.s., \; \forall i \in \mathbb{Z}_N. \tag{65}$$

Suppose for some $i \in \mathbb{Z}_N$, (65) does not hold, and let $A^i \triangleq \{(t,\omega) : g^i(t,\omega) < 0\}$. Since $g^i(t)$ is $\mathcal{G}^i_{0,t}-$measurable $\forall t \in [0,T]$ we can choose u^i in (63) as

$$u^i_t \triangleq \begin{cases} \nu \text{ on } A^i \\ u^{i,o}_t \text{ outside } A^i \end{cases}$$

together with $u^j_t = u^{j,o}_t, j \neq i, j \in \mathbb{Z}_N$. Substituting this in (63) we arrive at $\int_{A^i} g^i(t,\omega) ds d\mathbb{P} \geq 0$, which contradicts the definition of A^i, unless A^i has measure zero. Hence, (65) holds which is precisely (43). This completes Part **(I)**.

(II). By Theorem 4 the necessary conditions for team optimality satisfy those in Part **(I)** with $\mathcal{G}^i_{0,t}$ replaced by $\mathcal{G}^{z^{i,u}}_{0,t}$.

■

The following remark helps identifying the martingale term in the adjoint process.

Remark 2. *The arguments in the derivation of Theorem 6 involving the Riesz representation theorem for Hilbert space martingales, determine the martingale term of the adjoint process* $M_t = \int_0^t \psi_x^o(s)\sigma(s, x^o(s), u_s^o)dW(s)$*, dual to the first martingale term in the variational equation (23), provided* $\psi_x(\cdot)$ *exists (i.e.,* $f_{xx}, \sigma_{xx}, \ell_{xx}, \varphi_{xx}$ *exist and are uniformly bounded). Hence,* Q *in the adjoint equation (38), is identified as* $Q(t) \equiv \psi_x(t)\sigma(t, x(t), u_t)$*. When the diffusion term* $\sigma(\cdot,\cdot,\cdot)$ *is independent of* x*, given by* $\sigma(t,u)$*, then since* $\langle V_Q(t), \zeta\rangle = tr(Q^*(t)\sigma_x(t,x,u_t;\zeta))$ *we have* $V_Q(t) = 0, \forall t \in [0,T]$ *(e,g., the spatial derivative of the diffusion term is zero).*

It is interesting to note that the necessary conditions, for a $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ or $u^o \in \mathbb{U}_{rel}^{(N),z^u}[0,T]$ to be a person-by-person optimal policy, can be derived following similar steps as given in Theorem 6, and that these necessary conditions are the same as the necessary conditions for the team optimal strategy. This is stated as a Corollary.

Corollary 1. *(Necessary conditions for person-by-person optimality) Consider Problem 2 under Assumptions 2, 3. Under the conditions of Theorem 6, Part* **(I)***, for an element* $u^o \in \mathbb{U}_{rel}^{(N)}[0,T]$ *with the corresponding solution* $x^o \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega,\mathbb{R}^n))$ *to be a person-by-person optimal strategy, it is necessary that statements* **(1), (3)** *of Theorem 6, and Part* **I***, with statement* **(2)** *replaced by*

$$\mathbb{E}\int_0^T \mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{-i,o}, u_t^i - u_t^{i,o})dt \geq 0, \quad \forall u^i \in \mathbb{U}_{rel}^i[0,T], \ \ \forall i \in \mathbb{Z}_N. \tag{66}$$

hold. Similar conclusions hold for strategies $\mathbb{U}_{rel}^{(N),z^u}[0,T]$.

Proof: Primarily, the derivation is based on the same procedure as that of Theorem 6. The only difference is, that in this case, the variations of the DM policies are carried out in the direction of individual members while the rest of the members carry optimal policy.

■

Clearly, every team optimal strategy for Problem 1 is a person-by-person optimal strategy for Problem 2. Hence person-by-person optimality is weaker than team optimality. By comparing the statements of Theorem 6 and Corollary 1, it is clear that statements **(1)** and **(3)** coincide, while the only difference are the variational inequalities (42) and (66). However, (66) implies (42), and it can be shown that (42) implies (66). Indeed, if (66) is violated for some $j \in \mathbb{Z}_N$ then by choosing all other $u^i = u^{i,o}, \forall i \in \mathbb{Z}_N, i \neq j$, the right side of (42) will be negative,

which is a contradiction. This observation is new, and has not been documented in the static team game literature [23].

Remark 3. *From the above necessary conditions one can deduce the necessary conditions for full centralized information and partial centralized information. We state these conditions below.*

(1) Centralized Full Information Structures. *Consider Problem 1 under the conditions of Theorem 6, Part* **(I)**, *and assume u^i are adapted to $\mathbb{F}_T$, $\forall i \in \mathbb{Z}_N$. The necessary conditions are given by the following point wise almost sure inequalities*

$$\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), \mu) \geq \mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u^o_t),$$
$$\forall \mu \in \mathcal{M}_1(\mathbb{A}^{(N)}),\ a.e.\ t \in [0,T],\ \mathbb{P}-a.s., \tag{67}$$

where $\{x^o(t), \psi^o(t), Q^o(t) : t \in [0,T]\}$ are the solutions of the Hamiltonian system (40), (41), (38), (39). This corresponds to the classical case [8].

Moreover, if the strategies are based on centralized state feedback information, that is, u^i are adapted to the information $\mathcal{G}^{x^u}_T, \forall i \in \mathbb{Z}_N$, then under the conditions of Theorem 6, Part **(II)** *the previous optimality conditions are replaced by*

$$\mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), \mu)|\mathcal{G}^{x^o}_{0,t}\Big\} \geq \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), u^o_t)|\mathcal{G}^{x^o}_{0,t}\Big\},$$
$$\forall \mu \in \mathcal{M}_1(\mathbb{A}^{(N)}), a.e. t \in [0,T], \mathbb{P}|_{\mathcal{G}^{x^o}_{0,t}} - a.s. \tag{68}$$

(2) Centralized Partial Information Structures. *Consider Problem 1 under the conditions of Theorem 6, Part* **(I)** *and Part* **(II)** *and suppose that each u^i is adapted to the centralized partial information $\mathcal{G}_T \subset \mathbb{F}_T$, and $\mathcal{G}^{z^u}_T \subset \mathcal{F}^{x^u}_{0,T}$, respectively. Then the necessary condition is given by*

$$\mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi^o(t), Q^o(t), \mu)|\mathcal{K}_{0,t}\Big\} \geq \mathbb{E}\Big\{\mathbb{H}(t, x^o(t), \psi(t), Q(t), u^o_t)|\mathcal{K}_{0,t}\Big\},$$
$$\forall \mu \in \mathcal{M}_1(\mathbb{A}^{(N)}), a.e. t \in [0,T], \mathbb{P}|_{\mathcal{K}_{0,t}} - a.s. \tag{69}$$

where $\mathcal{K}_{0,t}$ is a sub-sigma algebra of any of the sigma algebras indicated above.

Finally, we mention two important results derived in [36] which have direct extensions to the current paper. The first addresses existence of measurable relaxed team optimal strategy associated with the minimization of the Hamiltonian, and the second addresses existence of realizable relaxed strategies by regular strategies.

B. Sufficient Conditions of Optimality

In this section, we show that the necessary conditions of optimality (43) are also sufficient under certain convexity conditions.

Theorem 7. *(Sufficient conditions for team optimality) Consider Problem 1 and suppose Assumptions 2, 3 hold. Under the conditions of Theorem 6, Part* **(I)**, *let* $(u^o(\cdot), x^o(\cdot))$ *denote any control-state pair (decision-state) and let* $\psi^o(\cdot)$ *the corresponding adjoint processes.*
Suppose the following conditions hold:

(C4) $\mathbb{H}(t, \cdot, \zeta, M, \nu), t \in [0, T]$ *is convex in* $\xi \in \mathbb{R}^n$*;*

(C5) $\varphi(\cdot)$ *is convex in* $\xi \in \mathbb{R}^n$*.*

Then $(u^o(\cdot), x^o(\cdot))$ *is team optimal if it satisfies (43). In other words, necessary conditions are also sufficient. For feedback strategies* $\mathbb{U}_{rel}^{(N),z^u}[0, T]$ *the same statement holds under the conditions of Theorem 6, Part* **(II)**.

Proof: We shall prove the sufficiency under the conditions of Theorem 6, **(I)**, that is, the admissible strategies $\mathbb{U}_{rel}^{(N)}[0, T]$, since the derivation is precisely the same for the case Part **(II)**. Let $u^o \in \mathbb{U}_{rel}^{(N)}[0, T]$ denote a candidate for the optimal team decision and $u \in \mathbb{U}_{rel}^{(N)}[0, T]$ any other decision. Then

$$J(u^o) - J(u) = \mathbb{E}\bigg\{ \int_0^T \Big(\ell(t, x^o(t), u_t^o) - \ell(t, x(t), u_t) \Big) dt + \Big(\varphi(x^o(T)) - \varphi(x(T)) \Big) \bigg\}. \quad (70)$$

By the convexity of $\varphi(\cdot)$ then

$$\varphi(x(T)) - \varphi(x^o(T)) \geq \langle \varphi_x(x^o(T)), x(T) - x^o(T) \rangle. \quad (71)$$

Substituting (71) into (70) yields

$$\begin{aligned} J(u^o) - J(u) \leq \mathbb{E}\Big\{ \langle \varphi_x(x^o(T)), x^o(T) - x(T) \rangle \Big\} \\ + \mathbb{E}\bigg\{ \int_0^T \Big(\ell(t, x^o(t), u_t^o) - \ell(t, x(t), u_t) \Big) dt \bigg\}. \end{aligned} \quad (72)$$

Applying the Ito differential rule to $\langle \psi^o, x - x^o \rangle$ on the interval $[0, T]$ and then taking expecation

we obtain the following equation.

$$\begin{aligned}
\mathbb{E}\Big\{\langle\psi^o(T),x(T)-x^o(T)\rangle\Big\} &= \mathbb{E}\Big\{\langle\psi^o(0),x(0)-x^o(0)\rangle\Big\} \\
&+\mathbb{E}\Big\{\int_0^T\langle -f_x^*(t,x^o(t),u_t^o)\psi^o(t)dt - V_{Q^o}(t) - \ell_x(t,x^o(t),u_t^o), x(t)-x^o(t)\rangle dt\Big\} \\
&+\mathbb{E}\Big\{\int_0^T\langle\psi^o(t), f(t,x(t),u_t)-f(t,x^o(t),u_t^o)\rangle dt\Big\} \\
&+\mathbb{E}\Big\{\int_0^T tr(Q^{*,o}(t)\sigma(t,x(t),u_t)-Q^{*,o}(t)\sigma(t,x^o(t),u_t^o))dt\Big\} \\
&= -\mathbb{E}\Big\{\int_0^T\langle\mathbb{H}_x(t,x^o(t),\psi^o(t),Q^o(t),u_t^o), x(t)-x^o(t)\rangle dt \\
&+\mathbb{E}\Big\{\int_0^T\langle\psi^o(t), f(t,x(t),u_t)-f(t,x^o(t),u_t^o)\rangle dt\Big\} \\
&+\mathbb{E}\Big\{\int_0^T tr(Q^{*,o}(t)\sigma(t,x(t),u_t)-Q^{*,o}(t)\sigma(t,x^o(t),u_t^o))dt\Big\} \qquad (73)
\end{aligned}$$

Note that $\psi^o(T)=\varphi_x(x^o(T))$. Substituting (73) into (72) we obtain

$$\begin{aligned}
J(u^o)-J(u) \leq &\mathbb{E}\Big\{\int_0^T\Big[\mathbb{H}(t,x^o(t),\psi^o(t),Q^o(t),u_t^o)-\mathbb{H}(t,x(t),\psi^o(t),Q^o(t),u_t)\Big]dt\Big\} \\
&-\mathbb{E}\Big\{\int_0^T\langle\mathbb{H}_x(t,x^o(t),\psi^o(t),Q^o(t),u_t^o), x^o(t)-x(t)\rangle dt\Big\}. \qquad (74)
\end{aligned}$$

Since by hypothesis $\mathbb{H}$ is convex in $\xi\in\mathbb{R}^n$ and linear in $\nu\in\mathcal{M}_1(\mathbb{A}^{(N)})$, $\mathbb{H}$ is convex in both $(\xi,\nu)\in\mathbb{R}^n\times\mathcal{M}_1(\mathbb{A}^{(N)})$. Using this fact in (74) we readily obtain

$$J(u^o)-J(u)\leq \mathbb{E}\int_0^T < \mathbb{H}(t,x^o(t),\psi^o(t),Q^o(t),\cdot), u_t^o(\cdot)-u_t(\cdot) > dt \leq 0, \quad \forall u\in\mathbb{U}_{rel}^{(N)}[0,T], \qquad (75)$$

where the last inequality follows from (43). This proves that u^o optimal and hence the necessary conditions are also sufficient.

■

Under conditions similar to those of Theorem 7, we can verify that a strategy is person-by-person optimal for Problem 2 if it satisfies (43); this is stated as a corollary. Indeed, the necessary conditions for team optimality and person-by-person optimality are equivalent, and person-by-person optimality implies team optimality.

Theorem 8. *(Sufficient conditions for person-by-person optimality) Consider Problem 2 and suppose Assumptions 2, 3 hold. Under the conditions of Theorem 6, Part* **(I)**, *let* $(u^o(\cdot),x^o(\cdot))$

denote any control-state pair and let $\psi^o(\cdot)$ *the corresponding adjoint processes.*
Suppose the conditions of Theorem 7, **(C4), (C5)** *hold.*
Then $(u^o(\cdot), x^o(\cdot))$ *is player-by-player optimal if it satisfies (43).*
For feedback strategies $\mathbb{U}_{rel}^{(N),z^u}[0,T]$ *the above statements hold under the conditions of Theorem 6, Part* **(II)***.*

Proof: The proof is similar to that of Theorem 7. ■

V. Optimality Conditions for Regular Strategies

In the development of the necessary and sufficient conditions of optimality given in the previous section we have given conditions which assert the existence of optimal decisions from the class of relaxed decisions $\mathbb{U}_{rel}^{(N)}[0,T]$ and $\mathbb{U}_{rel}^{(N),z^u}[0,T]$ in Theorem 1.
The main observation of this section is that, if optimal regular decisions exist from the admissible class $\mathbb{U}_{reg}^{(N)}[0,T] \subset \mathbb{U}_{rel}^{(N)}[0,T]$ (or the feedback class) then the necessary and sufficient conditions of Theorem 6 and Theorem 7 can be specialized to the class of decision strategies which are simply Dirac measures concentrated $\{u_t^o : t \in [0,T]\} \in \mathbb{U}_{reg}^{(N)}[0,T]$ or $\mathbb{U}_{reg}^{(N),z^u}[0,T]$. The important advantage of the theory of relaxed controls is that the necessary conditions of optimality for ordinary controls follow readily from those of relaxed controls without requiring differentiability of the Hamiltonian or equivalently the drift and the diffusion coefficients f, σ with respect to the control variables.
Thus we simply state the necessary and sufficient conditions of optimality for regular decentralized decision strategies which follow as a corollary of Theorem 6, 7 by simply specializing to regular decision strategies given by Dirac measures along the regular decision strategies leading to the following Hamiltonian

$$\mathcal{H} : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \times \mathbb{A}^{(N)} \longrightarrow \mathbb{R},$$

where

$$\mathcal{H}(t,\xi,\zeta,M,\nu) \stackrel{\triangle}{=} \langle f(t,\xi,\nu), \zeta\rangle + tr(M^*\sigma(t,\xi,\nu)) + \ell(t,\xi,\nu), \quad t \in [0,T]. \tag{76}$$

Theorem 9. *(Regular team optimality conditions) Consider Problem 1 under the Assumptions of Theorem 6 with decisions (or controls) from the regular class taking values in* $\mathbb{A}^i$*, a closed, bounded and convex subset of* $\mathbb{R}^{d_i}$*,* $\forall i \in \mathbb{Z}_N$*.*

(I) *Let $\mathbb{F}_T$ denote the filtration generated by $x(0)$ and the Brownian motion W.*

Necessary Conditions. *For an element $u^o \in \mathbb{U}_{reg}^{(N)}[0,T]$ with the corresponding solution $x^o \in B_{\mathbb{F}_T}^{\infty}([0,T], L^2(\Omega, \mathbb{R}^n))$ to be team optimal, it is necessary that the following hold.*

(1) *There exists a semi martingale $m^o \in \mathcal{SM}_0^2[0,T]$ with the intensity process $(\psi^o, Q^o) \in L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n) \times L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$.*

(2) *The variational inequality is satisfied:*

$$\sum_{i=1}^{N} \mathbb{E}\Big\{ \int_0^T \Big(\mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^i, u_t^{-i,o}) - \mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^o)\Big) dt \Big\} \geq 0, \quad \forall u \in \mathbb{U}_{reg}^{(N)}[0,T]. \tag{77}$$

(3) *The process $(\psi^o, Q^o) \in L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n) \times L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$ is a unique solution of the backward stochastic differential equation (38), (39), with $\mathbb{H}$ replaced by $\mathcal{H}$ such that $u^o \in \mathbb{U}_{reg}^{(N)}[0,T]$ satisfies the point wise almost sure inequalities with respect to the σ-algebras $\mathcal{G}_{0,t}^i \subset \mathbb{F}_{0,t}$, $t \in [0,T], i = 1, 2, \ldots, N :$*

$$\mathbb{E}\Big\{ \Big(\mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^i, u_t^{-i,o}) - \mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^o)\Big) | \mathcal{G}_{0,t}^i \Big\} \geq 0, \quad \forall u^i \in \mathbb{A}^i, a.e. t \in [0,T], \mathbb{P}|_{\mathcal{G}_{0,t}^i} - a.s., i = 1, 2, \ldots, N. \tag{78}$$

Sufficient Conditions. *Let $(u^o(\cdot), x^o(\cdot))$ denote an admissible decision and state pair and $\psi^o(\cdot)$ the corresponding adjoint processes.*

Suppose the conditions **(C4), (C5)** *holds and in addition*

(C6) $\mathcal{H}(t, \xi, \zeta, M, \cdot), t \in [0,T]$, *is convex in $u \in \mathbb{A}^{(N)}$;*

Then $(x^o(\cdot), u^o(\cdot))$ is optimal if it satisfies (78).

(II) *Suppose $\mathbb{F}_T$ is the filtration generated by $x(0)$ and the Brownian motion W, and Assumptions 5 hold with decision policies from the regular class. The necessary and sufficient conditions for a feedback policy $u^o \in \mathbb{U}_{reg}^{(N), z^u}[0,T]$ to be optimal are given by the statements under Part* **(I)** *with $\mathcal{G}_{0,t}^i$ replaced by $\mathcal{G}_{0,t}^{z^{i,u}}, \forall t \in [0,T]$.*

Proof: Follows from Theorem 6, 7 by simply replacing relaxed controls by Dirac measures concentrated at $\{u_t^o : t \in [0,T]\} \in \mathbb{U}_{reg}^{(N)}[0,T]$ or $\mathbb{U}_{reg}^{(N),z^u}[0,T]$.

■

Person-by-person optimality conditions for regular decision strategies follow from their relaxed counterparts, as discussed above. Therefore we simply state the results as a corollary.

Corollary 2. *(Person-by-person optimality) Consider Problem 2 under the conditions of Theorem 9. Then the necessary and sufficient conditions of Theorem 9 hold with the variational inequality (77) replaced by*

$$\mathbb{E}\Big\{ \int_0^T \Big(\mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^i, u_t^{-i,o}) \\ - \mathcal{H}(t, x^o(t), \psi^o(t), Q^o(t), u_t^{i,o}, u_t^{-i,o})\Big) dt \Big\} \geq 0, \quad \forall u^i \in \mathbb{U}_{reg}^i[0,T], \ \ \forall i \in \mathbb{Z}_N. \tag{79}$$

Similar conclusions hold for strategies $\mathbb{U}_{reg}^{z^{i,u}}[0,T]$.

Proof: Follows from Corollary 2 by simply replacing relaxed controls by Dirac measures concentrated at $\{u_t^o : t \in [0,T]\} \in \mathbb{U}_{reg}^{(N)}[0,T]$ or $\mathbb{U}_{reg}^{(N),z^u}[0,T]$.

■

The optimality conditions are derived based on the assumption that the filtration $\mathbb{F}_T$ is generated by the system Brownian motions $\{W(t) : t \in [0,T]\}$. When this condition does not hold the optimality conditions are slightly modified as discussed in the next remark.

Remark 4. *Suppose $\mathbb{F}_T$ is not generated by Brownian motions $\{W(t) : t \in [0,T]\}$ but stochastic integrals with respect to $W(\cdot)$ are $\mathbb{F}_T-$martingales. Then by invoking the variation of the semi martingale representation due to Kunita-Watanabe (for the derivation see [41]) we have the following. If $(i) : L^2(\Omega, \mathbb{F}, \mathbb{P})$ is separable and (ii): $\mathbb{F}_T$ is right continuous having left limits, then any square integrable $\mathbb{F}_T$ martingale has the decomposition*

$$m(t) = m(0) + \int_0^t v(s)ds + \int_0^t \Sigma(s)dW(s) + M(t), \qquad t \in [0,T], \tag{80}$$

for some $v \in L^2_{\mathbb{F}_T}([0,T], \mathbb{R}^n)$, $\Sigma \in L^2_{\mathbb{F}_T}([0,T], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n))$, $\mathbb{R}^n-$valued $\mathbb{F}_{0,0}-$measurable random variable $m(0)$ having finite second moment, and $\{M(t) : t \in [0,T]\}$ right continuous square integrable $\mathbb{F}_T$ martingale, which is orthogonal to $\{W(t) : t \in [0,T]\}$. This representation is unique. Further, the stochastic integrals $\int_0^t \Sigma(s)dW(s)$ and $\int_0^t \Gamma(s)dM(s)$ are orthogonal martingales for L_2 integrands. In this case the adjoint equation given by (38), (39) is replaced

by

$$d\psi(t) = -\,\mathbb{H}_x(t, x(t), \psi(t), Q(t), u_t)dt + Q(t)dW(t) + dM(t), \quad t \in [0, T) \tag{81}$$

$$\psi(T) = \varphi_x(x(T)). \tag{82}$$

In view of the results obtained, we confirm that there are no limitations in applying classical theory of optimization to decentralized systems. Rather, the challenge is in the implementation of the new variational Hamiltonians and the computation the optimal strategies for specific examples. In Part II [38] of this two-part paper, we shall apply these optimality conditions to investigate various linear and nonlinear distributed stochastic team games and obtain closed form expressions for the optimal strategies for some of them.

VI. Conclusions and Future Work

In this paper we have considered team games for distributed stochastic dynamical decision systems, with decentralized noiseless information patterns for each DM, under relaxed and deterministic strategies. Necessary and sufficient optimality conditions with respect to team optimality and person-by-person optimality criteria are derived, based on Stochastic Pontryagin's minimum principle, while we also discussed existence of the optimal strategies.

The methodology is very general, and applicable to many areas. However, several additional issues remain to be investigated. Below, we provide a short list.

(F1) For team games with regular strategies and non-convex action spaces $\mathbb{A}^i, i = 1, 2, \ldots, N$, if the diffusion coefficients depend on the decision variables then it is necessary to derive optimality conditions based on second-order variations. The methodology presented to derive the necessary conditions of optimality can be easily extended to cover this case as well.

(F2) The derivation of optimality conditions can be used in other type of games such as Nash-equilibrium games with decentralized information structures for each DM, and minimax games.

(F3) The optimality conditions can be extended to distributed stochastic dynamical decision systems driven by both continuous Brownian motion processes and jump processes, such as Lévy or Poisson jump processes, by following the procedure of centralized strategies in [36].

(F4) The optimality conditions can be applied to specific examples with decentralized noiseless information structures. Some of these are presented in the companion paper [38].

(F5) The methodology can be extended to cover decentralized partial (noisy) information structures.

REFERENCES

[1] W. Fleming and R. Rischel, *Deterministic and Stochastic Optimal Control*. Springer Verlag, 1975.

[2] R. J. Elliott, "The optimal control of stochastic system," *SIAM Journal on Control and Optimization*, vol. 15, no. 5, pp. 756–778, 1977.

[3] J. M. Bismut, "An introductory approach to duality in optimal stochastic control," *SIAM Review*, vol. 30, pp. 62–78, 1978.

[4] N. U. Ahmed and K. L. Teo, *Optimal Control of Distributed Parameter Systems*. Elsevier North Holland, New York, Oxford, 1981.

[5] R. J. Elliott, *Stochastic Calculus and Applications*. Springer-Verlag, 1982.

[6] R. J. Elliott and M. Kohlmann, "The second order minimum principle and adjoint process," *Stochastics & Stochastic Reports*, vol. 46, pp. 25–39, 1994.

[7] S. Peng, "A general stochastic maximum principle for optimal control problems,," *SIAM Journal on Control and Optimization*, vol. 28, no. 4, pp. 966–979, 1990.

[8] J. Yong and X. Y. Zhou, *Stochastic Controls, Hamiltonian Systems and HJB Equations*. Springer-Verlag, 1999.

[9] N. U. Ahmed, *Dynamic Systems and Control with Applications*. World Scientific, New Jersey London, Singapore Beijing Shanghai, Hong Kong, Taipei, Chenna, 2006.

[10] A. Bensoussan, "Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions," *Stochastics*, vol. 9, no. 169-222, 1983.

[11] ——, *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1982.

[12] C. D. Charalambous and J. L. Hibey, "Minimum principle for partially observable nonlinear risk-sensitive control problems using measure-valued decompositions," *Stochastics & Stochastic Reports*, pp. 247–288, 1996.

[13] N. U. Ahmed and C. D. Charalambous, "Minimax games for stochastic systems subject to relative entropy uncertainty: Applications to SDE's on Hilbert spaces," *Journal of Mathematics of Control, Signals and System*, vol. 19, pp. 197–216, 2007.

[14] N. U. Ahmed, "Optimal relaxed controls for systems governed by impulsive differential inclusions," *Nonlinear Functional Analysis & Applications*, vol. 10, no. 3, pp. 427–460, 2005.

[15] H. S. Witsenhausen, "A counter example in stochastic optimum control," *SIAM Journal on Control and Optimization*, vol. 6, no. 1, pp. 131–147, 1968.

[16] ——, "Separation of estimation and control for discrete time systems," in *Proceedings of the IEEE*, 1971, pp. 1557–1566.

[17] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems-part i," *IEEE Transactions on Automatic Control*, vol. 17, no. 1, pp. 15–22, 1972.

[18] B.-Z. Kurtaran and R. Sivan, "Linear-Quadratic-Gaussian control with one-step-delay sharing pattern," *IEEE Transactions on Automatic Control*, pp. 571–574, 1974.

[19] N. R. Sandell and M. Athans, "Solution of some nonclassical LQG stochastic decision problems," *IEEE Transactions on Automatic Control*, vol. 19, no. 2, pp. 108–116, 1974.

[20] B.-Z. Kurtaran, “A concice derivation of the LQG one-step-delay sharing problem solution,” *IEEE Transactions on Automatic Control*, vol. 20, no. 6, pp. 808–810, 1975.

[21] P. Varaiya and J. Walrand, “On delay sharing patterns,” *IEEE Transactions on Automatic Control*, vol. 23, no. 3, pp. 443–445, 1978.

[22] Y. Ho, “Team decision theory and information structures,” *Proceedings of IEEE*, vol. 68, pp. 644–655, 1980.

[23] J. Krainak, J. L. Speyer, and S. I. Marcus, “Static team problems-part I: Sufficient conditions and the exponential cost criterion,” *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 839–848, 1982.

[24] ——, “Static team problems-part II: Affine control laws, projections, algorithms, and the LEGT problem,” *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 848–859, 1982.

[25] R. Bansar and T. Basar, “Stochastic teams with nonclassical information revisited: When is an affine law optimal,” *IEEE Transactions on Automatic Control*, vol. 32, no. 6, pp. 554–559, 1987.

[26] P. R. Wall and J. H. van Schuppen, “A class of team problems with discrete action spaces: Optimality conditions based on multimodularity,” *SIAM Journal on Control and Optimization*, vol. 38, no. 3, pp. 875–892, 2000.

[27] B. Bamieh and P. Voulgaris, “A convex characterization of distributed control problems in spatially invariant systems with communication constraints,” *Systems and Control Letters*, vol. 54, no. 6, pp. 575–583, 2005.

[28] M. Aicardi, F. Davoli, and R. Minciardi, “Decentralized optimal control of markov chains with a common past information,” *IEEE Transactions on Automatic Control*, vol. 32, no. 11, pp. 1028–1031, 1987.

[29] A. Nayyar, A. Mahajan, and D. Teneketzis, “Optimal control strategies in delayed sharing information structures,” *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1606–1620, 2011.

[30] J. H. van Schuppen, “Control of distributed stochastic systems-introduction, problems, and approaches,” in *International Proceedings of the IFAC World Congress*, 2011.

[31] L. Lessard and S. Lall, “A state-space solution to the two-player optimal control problems,” in *Proceedings of 49th Annual Allerton Conference on Communication, Control and Computing*, 2011.

[32] A. Mahajan, N. Martins, M. Rotkowitz, and S. Yuksel, “Information structures in optimal decentralized control,” in *In Proceedings of the 51st Conference on Decision and Control*, 2011.

[33] J. Marschak, “Elements for a theory of teams,” *Management Science*, vol. 1, no. 2, 1955.

[34] R. Radner, “Team decision problems,” *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 857–881, 1962.

[35] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven: Yale University Pres, 1972.

[36] N. U. Ahmed and C. D. Charalambous, “Stochastic minimum principle for partially observed systems subject to continuous and jump diffusion processes and driven by relaxed controls,” *SIAM Journal on Control and Optimization*, 2012, submitted, June 2012.

[37] J. H. van Schuppen, O. Boutin, P. L. Kempker, J. Komenda, T. Masopust, N. Pambakian, and A. C. M. Ran, “Control of distributed systems: Tutorial and overview,” *European Journal on Control*, pp. 579–602, 2012.

[38] C. D. Charalambous and N. U. Ahmed, “Centralized versus decentralized team games of distributed stochastic differential decision systems with noiseless information structures-Part II: Applications,” *Preprint*, 2012, draft: October 2012.

[39] R. Liptser and A. Shiryayev, *Statistics of Random Processes Vol.1*. Springer-Verlag New York, 1977.

[40] A. Bensoussan, *Lecture on Stochastic Control, Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1982.

[41] C. Dellacherie and P. Meyer, *Probabilites et Potentiel*. Hermann, Paris, , ch. Chapitres I and IV.